

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:08:02 ; Search time 64.5137 Seconds
(without alignments)
1990.064 Million cell updates/sec

Title: US-10-631-958-2

Perfect score: 1717

Sequence: 1 PKHLVFNPGGKQGRKRI.....KCSRNFRLRIHTNQDQ 326

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1717	100.0	326	9 US-09-784-810A-11	Sequence 11, Appl
2	1717	100.0	326	10 US-09-969-896-2	Sequence 2, Appli
3	1717	100.0	326	16 US-10-631-958-2	Sequence 2, Appli
4	1717	100.0	326	18 US-10-876-281-11	Sequence 11, Appl
5	1717	100.0	471	9 US-09-784-810A-6	Sequence 6, Appli
6	1717	100.0	471	18 US-10-876-281-6	Sequence 6, Appli
7	1640.5	95.5	537	10 US-09-969-896-10	Sequence 10, Appl
8	1640.5	95.5	537	15 US-10-262-511-40	Sequence 40, Appl
9	1640.5	95.5	537	16 US-10-631-958-10	Sequence 10, Appl
10	1640.5	95.5	537	16 US-10-618-941-121	Sequence 121, App
11	1640.5	95.5	562	10 US-09-969-896-11	Sequence 11, Appl

12	1640.5	95.5	562	16	US-10-631-958-11	Sequence 11, Appl
13	1632.5	95.1	537	14	US-10-325-597A-2	Sequence 2, Appli
14	529	30.8	588	16	US-10-425-115-361076	Sequence 361076,
15	476.5	27.8	182	9	US-09-784-810A-29	Sequence 29, Appl
16	476.5	27.8	182	18	US-10-876-281-29	Sequence 29, Appl
17	452	26.3	575	16	US-10-425-115-270874	Sequence 270874,
18	335	19.5	484	15	US-10-437-963-133133	Sequence 133133,
19	334	19.5	245	15	US-10-425-114-53957	Sequence 53957, A
20	333	19.4	487	15	US-10-424-599-190238	Sequence 190238,
21	327	19.0	490	16	US-10-425-115-249553	Sequence 249553,
22	321.5	18.7	613	16	US-10-437-963-195284	Sequence 195284,
23	318	18.5	760	16	US-10-425-115-213665	Sequence 213665,
24	311.5	18.1	381	16	US-10-425-115-249555	Sequence 249555,
25	306	17.8	520	15	US-10-424-599-202261	Sequence 202261,
26	305.5	17.8	470	16	US-10-425-115-282512	Sequence 282512,
27	305.5	17.8	521	15	US-10-425-114-68423	Sequence 68423, A
28	300	17.5	446	16	US-10-437-963-197631	Sequence 197631,
29	293	17.1	677	16	US-10-425-115-264087	Sequence 264087,
30	289.5	16.9	373	16	US-10-767-701-44927	Sequence 44927, A
31	288	16.8	764	16	US-10-437-963-135389	Sequence 135389,
32	287.5	16.7	757	16	US-10-437-963-180379	Sequence 180379,
33	282	16.4	490	14	US-10-053-510-19	Sequence 19, Appl
34	282	16.4	490	15	US-10-348-052-19	Sequence 19, Appl
35	282	16.4	490	15	US-10-622-011-19	Sequence 19, Appl
36	282	16.4	641	15	US-10-348-052-28	Sequence 28, Appl
37	282	16.4	641	16	US-10-622-011-28	Sequence 28, Appl
38	273	15.9	144	9	US-09-784-810A-8	Sequence 8, Appli
39	273	15.9	144	18	US-10-876-281-8	Sequence 20, Appl
40	270.5	15.8	524	14	US-10-053-510-20	Sequence 20, Appl
41	270.5	15.8	524	15	US-10-348-052-20	Sequence 20, Appl
42	270.5	15.8	524	16	US-10-622-011-20	Sequence 20, Appl
43	270.5	15.8	907	15	US-10-348-052-29	Sequence 29, Appl
44	270.5	15.8	907	16	US-10-622-011-29	Sequence 29, Appl
45	269.5	15.7	552	16	US-10-437-963-161184	Sequence 161184,

ALIGNMENTS

RESULT 1

US-09-784-810A-11
; Sequence 11, Application US/09784810A
; Patent No. US20020082203A1

; GENERAL INFORMATION:

; APPLICANT: RASTELLI, LUCA

; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING

; FILE REFERENCE: 10716-08

; CURRENT APPLICATION NUMBER: US/09/784,810A

; CURRENT FILING DATE: 2001-02-14

; PRIOR APPLICATION NUMBER: 60/182,360

; PRIOR FILING DATE: 2000-02-14

; PRIOR APPLICATION NUMBER: 60/191,261

; PRIOR FILING DATE: 2000-03-22

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11

; LENGTH: 326

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-784-810A-11

Query Match 100.0%; Score 1717; DB 9; Length 326;
Best Local Similarity 100.0%; Pred. No. 2.8e-172;

Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLVFNPGGKQGRKRIYERKVPFLTLASITTDIIGNKFYNYVEVITEHANQAKE 60

DB 1 PKHLVFNPGGKQGRKRIYERKVPFLTLASITTDIIGNKFYNYVEVITEHANQAKE 60

QY 61 TLYEINIDKYGIVCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRVLPVPSLRIGIIPA 120

DB 61 TLYEINIDKYGIVCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRVLPVPSLRIGIIPA 120

```
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCGFYGDIK 180
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCGFYGDIK 180
Qy 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240
Db 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240
Qy 241 QOLEEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMNSCACRRSPRGLSPAHLGDGSS 300
Db 241 QOLEEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMNSCACRRSPRGLSPAHLGDGSS 300
Qy 301 DLILIRKSRNFNLFRLIRHTNQDQ 326
Db 301 DLILIRKSRNFNLFRLIRHTNQDQ 326
```

RESULT 2

```
US-09-969-896-2
; Sequence 2, Application US/09969896
; Publication No. US20030125533A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 2
; LENGTH: 326
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-969-896-2
```

```
Query Match 100.0%; Score 1717; DB 10; Length 326;
Best Local Similarity 100.0%; Pred. No. 2.8e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTTDIIGNKFYVNVYVEVITEHANOAKE 60
Db 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTTDIIGNKFYVNVYVEVITEHANOAKE 60
Qy 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
Db 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCGFYGDIK 180
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCGFYGDIK 180
Qy 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240
Db 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240
Qy 241 QOLEEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMNSCACRRSPRGLSPAHLGDGSS 300
Db 241 QOLEEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMNSCACRRSPRGLSPAHLGDGSS 300
Qy 301 DLILIRKSRNFNLFRLIRHTNQDQ 326
Db 301 DLILIRKSRNFNLFRLIRHTNQDQ 326
```

RESULT 3

```
US-10-631-958-2
; Sequence 2, Application US/10631958
```

```
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-958-2
```

```
Query Match 100.0%; Score 1717; DB 16; Length 326;
Best Local Similarity 100.0%; Pred. No. 2.8e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTTDIIGNKFYVNVYVEVITEHANOAKE 60
Db 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTTDIIGNKFYVNVYVEVITEHANOAKE 60
Qy 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
Db 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCGFYGDIK 180
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCGFYGDIK 180
Qy 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240
Db 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240
Qy 241 QOLEEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMNSCACRRSPRGLSPAHLGDGSS 300
Db 241 QOLEEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMNSCACRRSPRGLSPAHLGDGSS 300
Qy 301 DLILIRKSRNFNLFRLIRHTNQDQ 326
Db 301 DLILIRKSRNFNLFRLIRHTNQDQ 326
```

RESULT 4

```
US-10-876-281-11
; Sequence 11, Application US/10876281
; Publication No. US20050123942A1
```

```
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 326
; TYPE: PRT
```



```
; ORGANISM: Homo sapiens
US-10-876-281-11

Query Match      100.0%; Score 1717; DB 18; Length 326;
Best Local Similarity 100.0%; Pred. No. 2.8e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
DB 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 120
DB 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 120

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180

QY 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

QY 241 QOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
DB 241 QOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300

QY 301 DLILIRKCSRNFNRLRHTNQDDQ 326
DB 301 DLILIRKCSRNFNRLRHTNQDDQ 326

QY 301 DLILIRKCSRNFNRLRHTNQDDQ 326
DB 301 DLILIRKCSRNFNRLRHTNQDDQ 326

RESULT 5
US-09-784-810A-6
; Sequence 6, Application US/09784810A
; Publication No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-810A-6

Query Match      100.0%; Score 1717; DB 9; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
DB 53 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 112

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 120
DB 113 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 172

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 173 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 232

QY 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

Query Match      100.0%; Score 1717; DB 9; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
DB 53 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 112

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 120
DB 113 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 172

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 173 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 232

QY 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

; ORGANISM: Homo sapiens
US-09-784-810A-6

Query Match      100.0%; Score 1717; DB 18; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
DB 53 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 112

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 120
DB 113 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 172

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 173 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 232

QY 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

; ORGANISM: Homo sapiens
US-10-876-281-6
; Sequence 6, Application US/10876281
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-876-281-6

Query Match      100.0%; Score 1717; DB 18; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
DB 53 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 112

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 120
DB 113 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 172

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 173 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 232

QY 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

Query Match      100.0%; Score 1717; DB 18; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
DB 53 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 112

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 120
DB 113 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQVSAGVDQNHPRAVLPSSLRIGIIPA 172

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 173 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSLTLLRYSVSLGFGFYGDIK 232

QY 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

; ORGANISM: Homo sapiens
US-09-969-896-10
; Sequence 10, Application US/09969896
; Publication No. US2003012553A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; Kinase-Like Protein
```



```
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-958-10

Query Match      95.5%; Score 1640.5; DB 16; Length 537;
Best Local Similarity 96.3%; Pred. No. 6.9e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 1 PKHLLVFNPFGGKQGGKRIYERKVAPLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 60
DB 130 PKHLLVFNPFGGKQGGKRIYERKVAPLFTLASITTTDII-----VTEHANOAKE 178

QY 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120
DB 179 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 238

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHNSTLLRYSVSLLGYGYGDIK 180
DB 239 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHNSTLLRYSVSLLGYGYGDIK 298

QY 181 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCROSK 240
DB 299 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCROSK 358

QY 241 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
DB 359 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418

QY 301 DLILIRKSRFNLRLIRHTNQDQ 326
DB 419 DLILIRKSRFNLRLIRHTNQDQ 444

RESULT 11
US-09-969-896-11
; Sequence 11, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-896-11

Query Match      95.5%; Score 1640.5; DB 10; Length 562;
Best Local Similarity 96.3%; Pred. No. 7.4e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 1 PKHLLVFNPFGGKQGGKRIYERKVAPLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 60
DB 155 PKHLLVFNPFGGKQGGKRIYERKVAPLFTLASITTTDII-----VTEHANOAKE 203

QY 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120
DB 204 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 263

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHNSTLLRYSVSLLGYGYGDIK 180
DB 264 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHNSTLLRYSVSLLGYGYGDIK 323

QY 181 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCROSK 240
DB 324 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCROSK 383

QY 241 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
DB 384 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 443
```

```
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-958-10

Query Match      95.5%; Score 1640.5; DB 16; Length 537;
Best Local Similarity 96.3%; Pred. No. 6.9e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 1 PKHLLVFNPFGGKQGGKRIYERKVAPLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 60
DB 130 PKHLLVFNPFGGKQGGKRIYERKVAPLFTLASITTTDII-----VTEHANOAKE 178

QY 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120
DB 179 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 238

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHNSTLLRYSVSLLGYGYGDIK 180
DB 239 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHNSTLLRYSVSLLGYGYGDIK 298

QY 181 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCROSK 240
DB 299 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCROSK 358

QY 241 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
DB 359 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418

QY 301 DLILIRKSRFNLRLIRHTNQDQ 326
DB 419 DLILIRKSRFNLRLIRHTNQDQ 444

RESULT 10
US-10-618-941-121
; Sequence 121, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 121
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-941-121

Query Match      95.5%; Score 1640.5; DB 16; Length 537;
Best Local Similarity 96.3%; Pred. No. 6.9e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
```

Qy 301 DLILIRKCSRFNRLFLIRHTNQDQ 326
Db 444 DLILIRKCSRFNRLFLIRHTNQDQ 469

RESULT 12

US-10-631-958-11
; Sequence 11, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-958-11

Query Match 95.5%; Score 1640.5; DB 16; Length 562;
Best Local Similarity 96.3%; Pred. No. 7.4e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 1 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFFVYVVEVITEHANQAKE 60
Db 155 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANQAKE 203
Qy 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSLSRIGIIPA 120
Db 204 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSLSRIGIIPA 263
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIK 180
Db 264 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIK 323
Qy 181 DSEKRWGLARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240
Db 324 DSEKRWGLARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSK 383
Qy 241 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
Db 384 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 443
Qy 301 DLILIRKCSRFNRLFLIRHTNQDQ 326
Db 444 DLILIRKCSRFNRLFLIRHTNQDQ 469

RESULT 13

US-10-315-597A-2
; Sequence 2, Application US/10315597A
; Publication No. US20030162206A1
; GENERAL INFORMATION:
; APPLICANT: Sugiyura, Masako
; APPLICANT: Kono, Keita
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It
; FILE REFERENCE: 02658CIP/HG
; CURRENT APPLICATION NUMBER: US/10/315,597A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: JP 2000-178039
; PRIOR FILING DATE: 2000-06-14

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-315-597A-2

Query Match 95.1%; Score 1632.5; DB 14; Length 537;
Best Local Similarity 95.7%; Pred. No. 4.8e-163;
Matches 312; Conservative 2; Mismatches 1; Indels 11; Gaps 1;

Qy 1 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFFVYVVEVITEHANQAKE 60
Db 130 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANQAKE 178
Qy 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSLSRIGIIPA 120
Db 179 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSLSRIGIIPA 238
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIK 180
Db 239 GPTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIK 298
Qy 181 DSEKRWGLARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240
Db 299 DSEKRWGLARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSK 358
Qy 241 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
Db 359 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418
Qy 301 DLILIRKCSRFNRLFLIRHTNQDQ 326
Db 419 DLILIRKCSRFNRLFLIRHTNQDQ 444

RESULT 14

US-10-425-115-361076
; Sequence 361076, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 361076
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(588)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_92481C.1.pap
US-10-425-115-361076

Query Match 30.8%; Score 529; DB 16; Length 588;
Best Local Similarity 32.0%; Pred. No. 2.2e-46;
Matches 131; Conservative 57; Mismatches 107; Indels 114; Gaps 11;

Qy 1 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFFVYVVEVITEHANQAKE 60
Db 186 PKNLMVFVHPLCGKGRGVSNWE-TVYFLFAKAKTKVI-----VTERAGHAYD 233
Qy 61 T---LYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTOR----- 95

```
Db      234  TLSSLSDELKAFDGVAVGGDLGFLNEILLGILSRHKVSPPTPEGFGYVGSSDNCREQ 293
QY      96  -----SAGVQ-----NHPRVLVPSS-----112
Db      294  TRNWNFSKPTPDGNAVFLGNPSKDDQDPLSTSHPSSELEIPSNSTKEPCIGDQDNPV 353
QY     113  -----LRIGIIPAGSTDCVCTVGTSDAETSALHIVVGDSLAMDVSVHNS-----161
Db     354  SPFDWFRGIIIPSGSTDAILVSTTGERDPVTSAIIILGRKVSLEDAQVVRWKTSPAAR 413
QY     162  ---LLRYSVSLGCGFYGDIIKDSEKRWGLGARYDFSGLTKTFLSHHCVGTVSFLPAQH 218
Db     414  VLPTRYAASPAGYGFYGEVIRESEKYRMWGPARYDFSGTMVFLKHSYGAKVAFDTPY 473
QY     219  T---VGSPRD-----RKP-----CRAGCFVCRSQKQLEEBEQKALYGLEAAED 259
Db     474  THSLTASAQDDITGAQLQSRWKKPRKII CRTNCFVCKEASTSGQNPDDVADNSRTICEN 533
QY     260  VVEWQVCGKFLAINATNMSCACRSRGLSPAHLGDSDDLIIIRKC 308
Db     534  -QKWMSEGHFLSVGAATVSCRNERAPDGLVADAHLSDGFLHLLLRDC 581
```

RESULT 15

```
US-09-784-810A-29
; Sequence 29, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 80432911
US-09-784-810A-29
```

```
Query Match      27.8%; Score 476.5; DB 9; Length 182;
Best Local Similarity 88.6%; Pred. No. 1.4e-41;
Matches 93; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY      1  PKHLLVFNPFGGKGKRIYERKVAPLTLASITTDIIGNKFYVYVEVITEHANQAKE 60
Db      89  PKHLLVFNPFGGKGKRIYERKVAPLTLASITTDII-----VTEHANQAKE 137

QY     61  TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPR 105
Db     138  TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPR 182
```

Search completed: September 3, 2005, 04:32:26
Job time : 66.5137 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: September 6, 2005, 02:19:54 ; Search time 503.298 Seconds
(without alignments)
3834.378 Million cell updates/sec

Title: US-10-631-958-2
Perfect score: 1717
Sequence: 1 PKHLVFINPFGKCGQKRI.....KCSRFNFLRLIRHTNQDQ 326

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cpn2.1/USPTO.spool/US10631958/runat.02092005.165814.4497/app.query.fasta_1.1941
-DB=N_Geneseq.16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CGN 1.1436 @runat.02092005.165814.4497 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1717	100.0	979	6	ABL40822 Human sph
2	1717	100.0	1840	5	Aad14426 Human sph
3	1640.5	95.5	1740	8	Ada05679 Human NOV
4	1640.5	95.5	3975	8	Acc70838 Human sph
5	1640.5	95.5	4231	3	Aaa50510 Human sph

6	1640.5	95.5	4413	6	ABL40828	Abi40828 Human sph
7	1640.5	95.5	4429	12	ADJ96598	Adj96598 Human lip
8	1640.5	95.5	4445	13	ADP55247	Adp55247 Human PRO
9	1633.5	95.1	4432	8	ABA70921	Abx70921 Novel hum
10	1632.5	95.1	4463	6	ABA96945	Abx96945 Human cer
11	1595.5	92.9	1740	12	ADN62844	Adn62844 Human NOV
12	1414.5	82.4	4702	13	ADS10370	Adsl10370 Human the
13	1400.5	81.6	2241	5	AAS77728	Aas77728 DNA encod
14	1323.5	77.1	2186	5	AAS77730	Aas77730 DNA encod
15	1055	61.4	753	3	AACT6031	Aac76031 Human ORF
16	986	57.4	1570	5	AACT7731	Aas77731 DNA encod
17	811	47.2	474	6	ABL40823	Abi40823 Human sph
18	679	39.5	426	5	AAS77727	Aas77727 DNA encod
19	626	36.5	411	5	AAS77729	Aas77729 DNA encod
20	613	35.7	2173	4	ABL25705	Abi25705 Drosophil
21	613	35.7	10337	4	ABL25704	Abi25704 Drosophil
22	508.5	29.6	329	6	ABL40824	Abi40824 Human sph
23	476.5	27.8	547	3	AACT6592	Aac76592 Human ORF
24	442	25.7	1774	12	ADO88890	Aoq88890 Novel hum
25	414.5	24.1	564	8	ABT23453	Abt23453 Immune-re
26	389	22.7	15185	4	AAK65589	Aak65589 Human inm
27	386	22.5	15181	4	AAK65588	Aak65588 Human inm
28	302	17.6	167	6	ABL40825	Abi40825 Human sph
29	300.5	17.5	817	6	ABQ99499	Abq99499 Human cod
30	300.5	17.5	817	13	ADS11669	Adsl11669 Human the
31	282	16.4	2020	4	ABL03297	Abi03297 Drosophil
32	282	16.4	2629	9	ACF35859	Acf35859 D. melano
33	282	16.4	4020	4	ABL03296	Abi03296 Drosophil
34	273	15.9	522	5	AAD14427	Aad14427 Partial r
35	270.5	15.8	2609	9	ACF35860	Acf35860 D. melano
36	270.5	15.8	2830	4	ABL03237	Abi03237 Drosophil
37	269	15.7	572	13	ACNS7851	Acn57851 Cotton gy
38	260	15.1	153	6	ABL40826	Abi40826 Human sph
39	243	14.2	1857	6	ABK90199	Abk90199 cDNA enco
40	243	14.2	2380	5	AAS14817	Aas14817 Human cDN
41	243	14.2	2380	6	ABL59533	Abi59533 Human sph
42	243	14.2	2380	10	ADE85298	Ade85298 Farnesyl
43	243	14.2	2380	10	ADE38416	Ade38416 Human can
44	243	14.2	2380	12	ADQ15181	Adq15181 Human can
45	243	14.2	2422	5	AAS85331	Aas85331 DNA encod

ALIGNMENTS

RESULT 1
ABL40822
ID ABL40822 standard; cDNA; 979 BP.
XX
AC ABL40822;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human sphingosine kinase-like protein encoding cDNA.
KW Human sphingosine kinase-like protein; intracellular signalling; gene;
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 2..789
FT /*tag= a
FT /product= "sphingosine kinase-like protein"
FT /note= "start and stop codons are not indicated"
XX
PN WO200228906-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-EP011516.
XX
PR 06-OCT-2000; 2000US-0238005P.

PT associated disorder especially cancer, restenosis or ischemia in a human.
 XX Claim 8; Page 94-95; 107pp; English.

CC The present invention relates to sphingosine kinase (SphK) polypeptides
 CC and nucleic acids encoding them. SphK is useful for treating a SphK-
 CC associated disorder especially cancers such as leukaemia, lymphoma,
 CC ovarian, breast, lung, colon, testicular, stomach and skin,
 CC atherosclerosis, restenosis or ischaemia and cell proliferative disease
 CC or disorder associated with vascular diseases. SphK gene is used in gene
 CC therapy and antisense-therapy. Sphingolipids serving as signalling
 CC molecules, have recently emerged as regulators of cell growth,
 CC differentiation, diverse cell phenotypes and cell death. Activation of
 CC SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
 CC endothelial cells. The present sequence is human sphingosine kinase
 CC (SphK) cDNA

SQ Sequence 1840 BP; 459 A; 468 C; 503 G; 410 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.85e-179 Length: 1840
 Score: 1717.00 Matches: 326
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0

US-10-631-958-2 (1-326) x AAD14426 (1-1840)

Qy 1 ProlvsHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
 Db 265 CCAAGCAATTACTGGTATTATATACCCGTTTGGAGAAAGGACAAAGGCGGATA 324
 Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
 Db 325 TATGAACAAAGTGGCACCAGTGTTCACCTTAGCCTCATCACCTGACATCATCGT 384
 Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
 Db 385 AACAAATCTATGTTAACTATGTAGAGTAATTACTGAACATGCTTAATCAGGCCAAGGAG 444
 Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
 Db 445 ACTCTGTATGAGATTAAATAGACAAATACACGGCATCGTGTGTGGCGGAGATGGT 504
 Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
 Db 505 ATGTTACGGAGGTCTGCACGGTCTGATTTGGAGGACGACAGAGGCGCGGGGTCCAC 564
 Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
 Db 565 CAGAACCAACCCCGGGCTGTGCTGCTCCAGTAGCTCCGATTTGGAATCATTTCCCGCA 624
 Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
 Db 625 GGGTCACCGAGTGGGTGTGTACTCCACCGTGGGCACGACGACGAGAACTCTGGCG 684
 Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
 Db 685 CTGCATATCGTTGTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACACACAGC 744
 Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
 Db 745 ACACCTCTTCGCTACTCCGTGTCCCTGTGCTGGCTACGGCTTCTACGGGGACATCATCAAG 804
 Qy 181 AspSerGluLysValLeuArgTyrLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
 Db 805 GACATGTAGAGAAACCGTGGTGTGGCTCTGCCAGATACGATTTTCAGGTTTAAAGACC 864
 Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
 Db 865 TTCCTCTCCCACTGCTATGAGGGACAGTGTCTTCTTCTCCCTGCGACACACACCGGTG 924

Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
 Db 925 GGATCTCAAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGCAGGCAAGCAAG 984
 Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
 Db 985 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAGGACGTG 1044
 Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
 Db 1045 GAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGT 1104
 Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
 Db 1105 GCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCT 1164
 Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
 Db 1165 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1224
 Qy 321 ThrAsnGlnAspGln 326
 Db 1225 ACCAACCCAGCAGGACCAG 1242
 RESULT 3
 ADA05679
 ID ADA05679 standard; cDNA; 1740 BP.
 AC ADA05679;
 XX
 XX 06-NOV-2003 (first entry)
 DT
 XX
 DE Human NOV9a encoding cDNA SEQ ID NO:39.
 KW human; NOV9; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective;
 KW antiparkinsonian; antilipemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 76..1689
 FT /*tag= a
 FT /product= "NOV9a"
 FT
 PN WO2003029424-A2.
 XX
 XX 10-APR-2003.
 XX
 XX 02-OCT-2002; 2002WO-US031373.
 XX
 XX 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327435P.
 PR 05-OCT-2001; 2001US-0327449P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339286P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.

Db	172	ACTCTGTATGAGATTAAACATAGACAATAACGACGGCATCGTCTGTCTGCCGGAGATGGT	231
Qy	81	MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp	100
Db	232	ATGTTTCAGCCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGGAGCCTCGGGGT	291
Qy	101	GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIlelleProAla	120
Db	292	CAGAACCACCCCGGGCTGTGCTGGTCCCAGTAGCTCCGGATTTCGAATTCATTTCCCGCA	351
Qy	121	GlySerThrAspCysValCysTyrsrThrValGlyThrSerAspAlaGluThrSerAla	140
Db	352	GGGTCAACGGACTGCGTGTTACTCACCGTGGGACCCAGCGACGACGACCAACCTCGGG	411
Qy	141	LeuHisIleValValGlyAspSerLeuAlaMetAspValserSerValHisHisAsnSer	160
Db	412	CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACACAGC	471
Qy	161	ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIlelleLys	180
Db	472	ACACTCCTTCGCTACTCCGCTGCTCGGTGTACGGCTTCTACCGGGACATCATCAAG	531
Qy	181	AspSerGluIlylsYsargTrpLeuGlyIleAlaAqTYrAspPheSerGlyLeuIlyThr	200
Db	532	GACAGTGAGAAAGAACGGTGGTGGTCTTGCAGATACGACTTTTTCAGGTTTAAAGACC	591
Qy	201	PheLeuSerHisCysTyrGluGlyThrValserPheLeuProAlaGlnHisThrVal	220
Db	592	TTCTCTCCCCACCTGCTATGAGGGGACAGTGTCTTCTCTCCCTGCACACACACGGTG	651
Qy	221	GlySerProArgPheArglysProCysAsgAlaGlyCysPheValCysArgGlnSerLys	240
Db	652	GGATCTCCAAGGATAGGAAGCCCTCGCGGACAGGATGCTTTGTTTGCAGGCAAAGCAAG	711
Qy	241	GlnGlnLeuGluGluGlnIlylsYsAlaLeuTyrGlyLeuGluAlaGluAspVal	260
Db	712	CAGCAGCTGGAGGAGGAGGACAGAAAGACATGTATGGTTTGAAGCTCGGAGGACGTG	771
Qy	261	GluGluTrpGlnValValCysGlyIlylsPheLeuAlaIleAsnAlaThrAsnMetSerCys	280
Db	772	GAGGATGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGT	831
Qy	281	AlaCysAargArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer	300
Db	832	GCTTGTGCGCGGAGCCCAAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT	891
Qy	301	AspIlelleuIleahrglysCysSerArgPheAsnPheIeuahrgPheIleuIleAargHis	320
Db	892	GACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTTCTGAGATTTTCTCATCGGCAC	951
Qy	321	ThrAsnGlnGlnAspGln	326
Db	952	ACCAACCCAGCAGGACCCAG	969
RESULT 5			
AAA50510			
ID	AAA50510 standard; cDNA; 4231 BP.		
XX	AAA50510;		
XX	05-DEC-2000 (first entry)		
DT			
XX	Human sphingosine kinase C cDNA.		
DE			
XX	Sphingosine kinase C; SKC; human; drug screening; infection;		
KW	antiinflammatory; antiallergic; anticancer; inflammation; allergy;		
KX	cancer; therapy; diagnosis; ds.		
XX	Homo sapiens.		
OS			
XX	Key Location/Qualifiers		
FH	71..1453		
FT	CDS		
FT	/*tag= a		

XX WO200052173-A2.
 XX PN
 XX PD
 XX PD 08-SEP-2000.
 XX PF 02-MAR-2000; 2000WO-CA000223.
 XX PR 02-MAR-1999; 99US-0122516P.
 XX PA (ALLX) NPS ALLELIX CORP.
 XX PI Munroe D, Gupta A, Falzone GR;
 XX DR WPI; 2000-572185/53.
 XX DR P-PSDB; AAY96059.
 XX PT New human sphingosine kinase A, B and C polynucleotides and polypeptides
 XX PT useful in e.g. chromosome and gene mapping, and detecting inflammation or
 XX PT disease associated with abnormal levels of sphingosine kinase expression.
 XX PS Disclosure; Fig 7; 81pp; English.
 XX CC The present sequence is that of an isolated polynucleotide encoding human
 CC sphingosine kinase C (SKC, see AAY96059), an enzyme that phosphorylates
 CC sphingosine to form sphingosine 1-phosphate. The polynucleotide was
 CC isolated from an HeLa cDNA library by PCR amplification. The invention
 CC provides polynucleotides (see AAY50508-10) and polypeptides (see AAY96057
 CC -59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC.
 CC The polynucleotides may be used as hybridization probes, in the
 CC construction of PCR primers for chromosome and gene mapping, in the
 CC recombinant production of SKA, SKB and SKC, and in the generation of
 CC antisense DNA or RNA. They can be used to detect inflammation or disease
 CC associated with abnormal levels of SK expression, or to detect
 CC differences in gene sequence between normal and carrier or affected
 CC individuals. Host cells expressing SK can be used in drug screening.
 CC Human SK specific antibodies, inhibitors, ligands or their analogues are
 CC useful as bioactive agents to treat inflammation or disease including
 CC viral, bacterial or fungal infections, allergic responses, mechanical
 CC injury associated with trauma, hereditary diseases, lymphoma or
 CC carcinoma, and other conditions with activate the genes of kidney, lung,
 CC heart, lymphoid or tissues of the nervous system
 XX SQ Sequence 4231 BP; 1022 A; 1021 C; 1086 G; 1067 T; 0 U; 35 Other;

Alignment Scores:
 Pred. No.: 4,63e-170 Length: 4231
 Score: 1640.50 Matches: 314
 Percent Similarity: 96.63% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 95.54% Indels: 11
 DB: 3 Gaps: 1

US-10-631-958-2 (1-326) x AAY50510 (1-4231)

Qy 1 ProLysHisLeuLeuValPheileAsnProPheGlyGlyGlyGlnGlyLysArgile 20
 Db 227 CCARAGCATTTACTGTTATTTATCAACCGTTTGGAGGAAAGGACAGGCAAGCGGATA 286
 Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleleGly 40
 Db 287 TATGAAAGAAAGTGGCACCACCTGTTCACCTTAGCCTCCTCCTCCTCCTCCTCCTCCT 343
 Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
 Db 344 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 373
 Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
 Db 374 ACTCTGTATGAGATTAAATACACAAATACACGCGCATCGTCTGTCTGGCGGAGATGTT 433
 Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
 Db 434 ATGTTTCAGCGGTGCTGCACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGGTCGAC 493

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
 Db 494 CAGAACCAACCCCGGGCTGTGCTGCCAGTAGACCTCGGATTGGATCATTCGCGCA 553
 Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
 Db 554 GGGTCAACGGGACTGCGTGTGTACTCCACCGTGGGCACACGACGACGAGAACCTCGCG 613
 Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSer 160
 Db 614 CTGCATATCGTTGTGGGACTCGCTGCCATGATGTCTCTCAGTCCACCAACAGC 673
 Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
 Db 674 ACATCTCTTCTGCTACTCGGTGCTCTGGCTACGGCTTCTACGGGGACATCATCAAG 733
 Qy 181 AspSerGluIlyLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
 Db 734 GACAGTGAGAAGAAACGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 793
 Qy 201 PheLeuSerHisHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThrVal 220
 Db 794 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTCTCTGACACACACCGGTG 853
 Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
 Db 854 GGATCTCAAGGGATAGAGAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAGCAAG 913
 Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
 Db 914 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGTTTGGAACTGCGGAGCGTG 973
 Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
 Db 974 GAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGT 1033
 Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
 Db 1034 GCTTGTGCGGGAGCCCGGAGGCTCTCCCGGCTGCCCATCTGGGAGAGCGGTCTTCT 1093
 Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
 Db 1094 GACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1153
 Qy 321 ThrAsnGlnAspGln 326
 Db 1154 ACCAACCCAGCAGGACCAG 1171
 RESULT 6
 ABL40828
 ID ABL40828 standard; cDNA; 4413 BP.
 AC ABL40828;
 XX 03-JUL-2002 (first entry)
 DE Human sphingosine kinase-like protein encoding cDNA.
 XX Human sphingosine kinase-like protein; intracellular signalling; gene;
 KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
 KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 CDS 1..1689
 FT /*tag= a
 FT /product= "sphingosine kinase-like protein"
 FT /note= "see ABB07857"
 FT 76..1689
 FT /*tag= b
 FT /product= "sphingosine kinase-like protein"


```

FT      /tag= d
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(4272,g)
FT      /*tag= e
FT      /standard_name= "Single nucleotide polymorphism"
XX
XX      WO2004006838-A2.
XX
XX      22-JAN-2004.
XX
XX      15-JUL-2003; 2003WO-US021730.
XX
XX      15-JUL-2002; 2002US-0395632P.
XX      (SUGEN-) SUGEN INC.
XX
XX      Whyte D, Manning G, Caenepeel S;
XX      WPI; 2004-122753/12.
XX      P-PSDB; ADJ96664.
XX
XX      New nucleic acid molecule encoding a kinase polypeptide, useful for
XX      preparing a composition for treating diseases or disorders, e.g., cancer,
XX      or neurological, immunological or inflammatory disorders.
XX
XX      Example 1; SEQ ID NO 55; 366pp; English.
XX
XX      This invention relates to a novel isolated, enriched or purified nucleic
XX      acid molecule that encodes a kinase polypeptide. Specifically, it relates
XX      to human tyrosine and serine/threonine protein kinases (PTK's and STRK's),
XX      as well as protein kinase-like enzymes. The present invention describes
XX      screening methods to identify agonists, antagonists and antibodies that
XX      can be used to modulate the activity or function of the mammalian kinase
XX      enzymes. As such, these compositions can be used for gene therapy
XX      purposes to treat diseases or disorders including cancer, immune-related
XX      diseases, cardiovascular disease, brain or neuronal associated disease,
XX      metabolic and inflammatory disorders. Accordingly, they exhibit
XX      cytosolic, neuroprotective, immunomodulator and antiinflammatory
XX      activities. This polynucleotide sequence is a human kinase DNA sequence
XX      of the invention.
XX
XX      SQ      Sequence 4429 BP; 1020 A; 1115 C; 1213 G; 1081 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      4.95e-170      Length:      4429
Score:          1640.50      Matches:      314
Percent Similarity: 96.63%      Conservative: 1
Best Local Similarity: 96.32%      Mismatches: 0
Query Match:      95.54%      Indels:      11
DB:              12      Gaps:      1

US-10-631-958-2 (1-326) x ADJ96598 (1-4429)
Qy      1      ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyeArgIle 20
Db      479      CCAAGCATTTACTGGTATTTATCAACCGTTTGGAGGAAAAGGACAAAGCGGATA 538
Qy      21      TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db      539      TATGAAGAAAAGTGGCCACCTGTTCACCTTAGCCCTCCATCACCACCTGACATCATC--- 595
Qy      41      AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db      596      -----GTTACTGAACATGCTTAATCAGGCCCAAGGAG 625
Qy      61      ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db      626      ACTCTGTATGAGATTAAACATACAAATACACGCGCATCGTCTGTGTCGGCGGAGATGGT 685
Qy      81      MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db      686      ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGAGGACGACGAGGAGCGCGGGTTCGAC 745

```

```

Qy      101      GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db      746      CAGAACACCCCGGGCTGTGCTGCCAGTAGCTCCCGATTGGAATCATTTCCGCA 805
Qy      121      GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db      806      GGGTCAACGAGCTGCGTGTATTCTCCACCGTGGGCACGACGACGAGAAACCTCGGCG 865
Qy      141      LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db      866      CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACAGC 925
Qy      161      ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db      926      ACCTCTTCGCTACTCGGTGCTCCCTGCTGGCTACGGCTTCTACGGGGACATCATCAAG 985
Qy      181      AspSerGluLysLysArgTTPLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db      986      GACAGTGAGAGAAACCGTGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGACC 1045
Qy      201      PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db      1046      TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACGGTG 1105
Qy      221      GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db      1106      GGATCTCCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGCAAGCAAG 1165
Qy      241      GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db      1166      CAGCAGCTGGAGGAGGACAGAAAGACACTGTATGTTGGAGCTTGGAGCTGGGAGCGTG 1225
Qy      261      GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db      1226      GAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGCCATCATGTCACCAACATCTCTGT 1285
Qy      281      AlaCysArgArgSerProArgGlyLeuSerProAlaHisLeuGlyAspGlySerSer 300
Db      1286      GCTTGTCCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 1345
Qy      301      AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db      1346      GACTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCAC 1405
Qy      321      ThrAsnGlnGlnAspGln 326
Db      1406      ACCAACCCAGCAGGACCAAG 1423
XX      ADP55247;
XX      ADP55247;
XX      18-NOV-2004 (first entry)
XX      Human PRO cDNA sequence SEQ ID NO:1223.
XX      human; PRO; immune related disease; inflammatory immune response;
XX      immune response stimulant; antiallergic; antianemic; antiarthritic;
XX      antirheumatic; antidiabetic; antiinflammatory; antipsoriatic;
XX      antitubercular; antithyroid; CNS; dermatological; gastrointestinal;
XX      haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
XX      nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
XX      virucide; gene therapy; gene; ss.
XX      Homo sapiens.
XX      OS
XX      PN      WO2004039956-A2.
XX      PD      13-MAY-2004.
XX      PF      28-OCT-2003; 2003WO-US034381.

```

XX 29-OCT-2002; 2002US-0422472P.
 XX (GETH) GENENTECH INC.
 XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
 XX Wood WI, Wu TD;
 XX WPI: 2004-376182/35.
 XX P-PSDB; ADP55248.
 XX
 XX New PRO polynucleotides and polypeptides, useful in diagnosing
 XX and treating an immune related disease, e.g. systemic lupus
 XX erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
 XX stimulating an immune response.
 XX
 XX Claim 2; SEQ ID NO 1223; 3009bp; English.
 XX
 XX The present invention describes an isolated PRO nucleic acid (I). Also
 XX described: (1) a vector comprising (1); (2) a host cell comprising the
 XX vector of (1); (3) a process for producing a PRO polypeptides; (4) an
 XX isolated PRO polypeptide; (5) a chimeric molecule comprising the
 XX polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
 XX antibody which specifically binds to a polypeptide of (4); (7) a
 XX composition of matter comprising a polypeptide of (4), an agonist or
 XX antagonist of the polypeptide or an antibody that binds to the
 XX polypeptide in combination with a carrier; (8) an article of manufacture
 XX comprising a container, a label on the container and a composition of
 XX matter of (7); (9) a method of treating an immune related disease in a
 XX mammal; (10) a method for determining the presence of a PRO polypeptide
 XX in a sample suspected of having the polypeptide; (11) a method of
 XX diagnosing an immune related disease or an inflammatory immune response
 XX in mammal; (12) a method of identifying a compound that inhibits or
 XX mimics the activity of or expression of a gene encoding a PRO polypeptide
 XX ; and (13) a method of stimulating the immune response in a mammal. The
 XX PRO sequences have anti-allergic, anti-nausea, anti-arthritis,
 XX anti-asthmatic, anti-diabetic, anti-inflammatory, anti-psoriasis,
 XX anti-rheumatic, antithyroid, CNS, dermatological, gastrointestinal,
 XX haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
 XX nephrotropic, neuroprotective, osteoporotic, respiratory, vasotropic and
 XX virucide activities, and can be used in gene therapy. The nucleic acid
 XX (I) and the encoded polypeptides, compositions, kits and methods are
 XX useful in diagnosing and treating an immune related disease and in
 XX stimulating an immune response. The present sequence represents a human
 XX PRO nucleotide sequence from the present invention.
 XX
 XX Sequence 4445 BP; 1018 A; 1134 C; 1216 G; 1077 T; 0 U; 0 Other;
 XX
 XX Alignment Scores:
 XX Pred. No.: 4,97e-170 Length: 4445
 XX Score: 1640.50 Matches: 314
 XX Percent Similarity: 96.63% Conservative: 1
 XX Best Local Similarity: 96.32% Mismatches: 0
 XX Query Match: 95.54% Indels: 11
 XX DB: 13 Gaps: 1
 XX
 XX US-10-631-958-2 (1-326) x ADP55247 (1-4445)
 XX
 XX 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
 XX 511 CCNAGGCAATTACTGTTATATATCAACCCGTTTGGAGAAAGGACAGGCGGAGATA 570
 XX
 XX 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
 XX 571 TATGAAAGAAAGTGGCCACCACTGTTCACTTAGCTTCATCCACCTGACATCATC--- 627
 XX
 XX 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
 XX 628 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 657
 XX
 XX 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
 XX 658 ACTCTGTATGAGATTAACTAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGT 717

QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
 DB 718 ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGGCGCGGGTTCGAC 777
 QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
 DB 778 CAGAACCAACCCCGGGCTGTGCTGCCAGTAGCTCCGATTCGGATTTGGAAATCATTTCCCGCA 837
 QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
 DB 838 GGGTCAACGGACTGCTGTGTTACTCCACCGTGGGCACGACGACGAGAACCTCGCGG 897
 QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
 DB 898 CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACACAGC 957
 QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
 DB 958 ACATCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCCCTCCCTGCACACACACGCGTG 1017
 QY 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
 DB 1018 GACAGTGAGAGAAACGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 1077
 QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
 DB 1078 TTCTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCCCTCCCTGCACACACACGCGTG 1137
 QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
 DB 1138 GGATCTCAAGGGATAGGAAGCCCTGCCGGGCAGATGCTTTGTTTGCAGGCAAGCAAG 1197
 QY 241 GlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
 DB 1198 CAGCAGCTGGAGGAGGAGGACAGAGAACACTGTATGTTTGGAAAGCTGCGGAGGACGCTG 1257
 QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
 DB 1258 GAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACCAACATGTCTCTGT 1317
 QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
 DB 1318 GCTTGTGCCGGAGGCCCGGCGGCTCTCCCGGCTGCCACTTGGGAGAGCGGGTCTTCT 1377
 QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
 DB 1378 GACCTCATCTCATCCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1437
 QY 321 ThrAsnGlnGlnAspGln 326
 DB 1438 ACCAACCAAGCAGGACGACCAG 1455
 XX
 XX RESULT 9
 XX ABX70921
 XX ID ABX70921 standard; cDNA; 4432 BP.
 XX AC ABX70921;
 XX XX
 XX 05-MAR-2003 (first entry)
 XX XX
 XX DE Novel human cDNA sequence #146.
 XX XX
 XX Human; gene; ss; nervous system disorder; peripheral neuropathy;
 XX Huntington's disease; amyotrophic lateral sclerosis; haemophilia;
 XX neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
 XX autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
 XX insulin-dependent diabetes mellitus; anaemia; thrombocytopenia; wound;
 XX ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;
 XX fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;
 XX coagulation disorder; cancer; tumour; inflammatory disease; septic shock;
 XX Crohn's disease; anaphylaxis; proliferation; chemotactic; chemokinetic;
 XX differential; stem cell growth factor; haematopoiesis; chemokinetic;

haemostatic; antiinflammatory; expressed sequence tag; EST.

KW	XX	OS	Homo sapiens.	Db	1708	-----GTTACTGAACATGCTAATACAGGAG	1737
XX	XX	XX	WO200281731-A2.	Qy	61	ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly	80
XX	XX	XX	17-OCT-2002.	Db	1738	ACTCTGATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGGT	1797
XX	XX	XX	29-JAN-2002; 2002WO-US001222.	Qy	81	MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp	100
XX	XX	XX	30-JAN-2001; 2001US-00774528.	Db	1798	ATGTTACAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGGTGCAC	1857
XX	XX	XX	(HYSE-) HYSEQ INC.	Qy	101	GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla	120
XX	XX	XX	(GOOD/) GOODRICH R. W.	Db	1858	CAGAACCAACCCCGGGCTGCTGCTCCAGTAGCTCCGAGATTGGAATCATTTCCGCA	1917
XX	XX	XX	Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;	Qy	121	GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla	140
XX	XX	XX	Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;	Db	1918	GGGTCAACGGACTGCGTGTGTACTCCACCGTGGGACCGACGACGAGAAACCTCGCG	1977
XX	XX	XX	WPI; 2003-058563/05.	Qy	141	LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer	160
XX	XX	XX	Novel polypeptide useful for treating neurodegenerative diseases, myeloid	Db	1978	CTGCATATCGTTGTTGGGACTCGCTGCCCATGGATGTGCTCCTCAGTCCACCAACAGC	2037
XX	XX	XX	or lymphoid cell disorders, bone disorders, mechanical and traumatic	Qy	161	ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleIys	180
XX	XX	XX	disorders, coagulation disorders, and inflammatory diseases.	Db	2038	ACACTCTTCGCTACTCGTGTCTCGTGGCTTACGGGCTTCTACGGGGACATCATCAAG	2097
XX	XX	XX	Claim 1; Page: 612pp; English.	Qy	181	AspSerGluLysLysValArgTyrLeuGluAlaArgTyrAspPheSerGlyLeuLysThr	200
XX	XX	XX	This invention relates to the cDNA sequences encoding an isolated novel	Db	2098	GACAGTGAGAAGAACCGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC	2157
XX	XX	XX	human polypeptide. The protein encoded by the nucleic acid of the	Qy	201	PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal	220
XX	XX	XX	invention is useful for treating central and peripheral nervous system	Db	2158	TTCTCTCTCCCACTGCTATGGAGGGACAGTGTCTTCTCTCCCTGCACACACACGGTG	2217
XX	XX	XX	diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic	Qy	221	GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIys	240
XX	XX	XX	lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,	Db	2218	GGATCTCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGTCAGGCAAGCAAG	2277
XX	XX	XX	Alzheimer's disease); autoimmune disease (e.g. systemic lupus	Qy	241	GlnGlnLeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal	260
XX	XX	XX	erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)	Db	2278	CAGCAGCTGGAGGAGGAGCAGAGAAAGACATGTATGTTTGGAACTGCGGAGGACGTG	2337
XX	XX	XX	; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopaenia)	Qy	261	GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys	280
XX	XX	XX	; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,	Db	2338	GAGGAGTGGCAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGT	2397
XX	XX	XX	osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head	Qy	281	AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer	300
XX	XX	XX	trauma); lung or liver fibrosis; reperfusion injury in various tissues;	Db	2398	GCTTGTGCGCGGAGCCCGGCGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCT	2457
XX	XX	XX	bacterial, viral or fungal infections; allergic conditions such as	Qy	301	AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis	320
XX	XX	XX	allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);	Db	2458	GACCTCATCTCTATCCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC	2517
XX	XX	XX	cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's	Qy	321	ThrAsnGlnGlnAspGln 326	
XX	XX	XX	disease, anaphylaxis). The protein may be used to inhibit the growth,	Db	2518	ACCAACCAAGCAGGACCCAG 2535	
XX	XX	XX	infection or function of infectious agents such as bacteria, fungi,	RESULT 10			
XX	XX	XX	cycles of rhythms. The protein may also have	ABA96945			
XX	XX	XX	proliferation/differentiation, stem cell growth factor, haematopoiesis	ID	ABA96945		
XX	XX	XX	regulation, immune stimulating or suppressing, chemotactic/chemokinetic,	XX	ABA96945		
XX	XX	XX	haemostatic and thrombolytic, receptor/ligand, and antiinflammatory	XX	20-MAY-2002 (first entry)		
XX	XX	XX	activities. The cDNA sequences of the invention are useful for expressing	XX	Human ceramide kinase hCERK1-encoding cDNA.		
XX	XX	XX	recombinant protein for analysis. The present sequence represents a novel	XX	Human; ceramide kinase; hCERK1; drug screening; gene therapy;		
XX	XX	XX	human cDNA sequence of the invention, this sequence is an expressed	XX	neurological disease; inflammation; human immunodeficiency virus;		
XX	XX	XX	sequence tag (EST) and was identified using subtractive hybridisation	XX	HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;		
XX	XX	XX		XX	cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;		
SQ			Sequence 4432 BP; 970 A; 1122 C; 1282 G; 1058 T; 0 U; 0 Other;				

Alignment Scores:
 Pred. No.: 2,94e-169 Length: 4432
 Score: 1633.50 MatChes: 313
 Percent Similarity: 96.32% Conservative: 1
 Best Local Similarity: 96.01% Mismatches: 1
 Query Match: 95.14% Indels: 11
 DB: 8 Gaps: 1

US-10-631-958-2 (1-326) x ABX70921 (1-4432)

Qy	1	ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db	1591	CCAAAGCAATTACTGGTATTTATCAACCCGTTTGAGGAAAGGCAAGGCAAGCGGATA 1650
Qy	21	TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db	1651	TATGAAGAAGAAAGTGGCCACTGTTCACCTTAGCTTCATCACCACCTGACATCATC--- 1707
Qy	41	AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60

KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;
 XX gene; ss.

OS Homo sapiens.

XX Location/Qualifiers

XX Key 124..1737

XX CDS /*tag= a

XX /product= "Human ceramide kinase hCERK1"

XX WO200196575-A1.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001WO-JP004889.

XX 14-JUN-2000; 2000JP-00178039.

XX (SANY) SANKYO CO LTD.

XX Sugiura M, Kono K, Kohama T;

XX WPI; 2002-179513/23.

XX P-PSDB; AAM49115.

XX Human ceramide kinase gene and the enzyme encoded by it for screening
 PT substances as drugs for neurological, inflammatory and other disorders.

XX Claim 5; Page 46-53; 61pp; Japanese.

CC This sequence represents cDNA encoding a human ceramide kinase designated
 CC hCERK1. The invention relates to hCERK1, nucleic acids encoding it,
 CC expression vectors and host cells containing hCERK1 nucleic acids, the
 CC recombinant production of hCERK1 and antibodies specific for hCERK1. The
 CC invention also encompasses methods of isolating hCERK1 from samples, the
 CC use of hCERK1 in drug screening, and the use of hCERK1 nucleic acid
 CC sequences in gene therapy. hCERK1 mediates the ATP-dependent 1-
 CC phosphorylation of ceramides and can be used to screen for therapeutic
 CC and preventive agents for a wide range of disorders. Such disorders
 CC include neurological disease, inflammation, human immunodeficiency virus
 CC (HIV) infection, type 2 diabetes, obesity, sepsis, arteriosclerosis and
 CC cancer

SQ Sequence 4463 BP; 1026 A; 1135 C; 1217 G; 1084 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 3.83e-169 Length: 4463
 Score: 1632.50 Matches: 312
 Percent Similarity: 96.32% Conservative: 2
 Best Local Similarity: 95.71% Mismatches: 1
 Query Match: 95.08% Indels: 11
 DB: 6 Gaps: 1

US-10-631-958-2 (1-326) x ABA96945 (1-4463)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLeuArgIle 20
 Db 511 CCAAGCATTCTACTGTTATTTATCAACCGTTTGGAGAAAAGGACAAAGGCGGATA 570
 Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
 Db 571 TATGAAGAAGAAGTGGACCACTGTTCACCTTAGCCTCCATCACCACCTGACATCATC--- 627
 Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
 Db 628 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 657
 Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysGly 80
 Db 658 ACTCTGTATGAGATTAAACATAGACAAATACGACGCGCATCGTGTGTCGGCGGAGATGT 717
 Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100

Db 718 ATGTTACGAGGAGTCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGTTCGAC 777
 Qy 101 GlnAenHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
 Db 778 CAGAACCAACCCCGGGCTGTGCTGGTCCCGAGTAGCTCCCGATTGGAATCATTTCCCGCA 837
 Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
 Db 838 GGTCAACCGGACTGCGTGTGTTACTCCACCGTGGGCACGACGACGACGACAACTTCGGCG 897
 Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
 Db 898 CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTAGTCCACCAACACAGC 957
 Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
 Db 958 ACATCTCTTCGCTACTCGCTGTCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 1017
 Qy 181 AspSerGluLysLysArgTyrPheLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
 Db 1018 GACAGTGAGAAAGAACCGGTGGTCTTCCACGATACGACTTTTTCAGGTTTAAAGACC 1077
 Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
 Db 1078 TTCTCTCTCCCACTGCTATGAAGGGACAGTGTCTTCTCTCCCTGCAACACACACGGTG 1137
 Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
 Db 1138 GGATCTCCAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGGAGCAAGCAAG 1197
 Qy 241 GlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
 Db 1198 CAGCAGCTGGAGGAGGAGCAGAGAAGACACTGTATGTTTGGAAAGCTGCGGAGGACGTG 1257
 Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
 Db 1258 GAGGAGTGGCAAGTGTCTGTGGGAAGTTCTGGCCATCAATGCCAACAAACATGTCCTGT 1317
 Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
 Db 1318 GCTTGTGCGGAGGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTCT 1377
 Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
 Db 1378 GACCTCATCTCATCCGGAATGCTCCAAGTTCAATTTCTGAGATTTCTCATCAGGCAC 1437
 Qy 321 ThrAsnGlnAspGln 326
 Db 1438 ACCAACCCAGCAGGACCAG 1455
 RESULT 11
 ID ADN62844 standard; DNA; 1740 BP.
 XX ADN62844
 AC ADN62844;
 XX 01-JUL-2004 (first entry)
 DE Human NOV9a DNA.
 XX ds; gene; human; NOVX; metabolic disorder; diabetes; obesity;
 KW infectious disease; anorexia; cancer; cancer-associated cachexia;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; hematopoietic disorder; dyslipidaemia;
 KW metabolic syndrome X; wasting disorder.
 XX Homo sapiens.
 OS US2004038223-A1.
 XX 26-FEB-2004.
 PD 01-OCT-2002; 2002US-00262511.
 PF

```

XX 02-OCT-2001; 2001US-0326483P.
PR 05-OCT-2001; 2001US-0327435P.
PR 05-OCT-2001; 2001US-0327449P.
PR 09-OCT-2001; 2001US-0327917P.
PR 09-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328044P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328849P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 18-OCT-2001; 2001US-0330309P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0339266P.
PR 24-OCT-2001; 2001US-0343629P.
PR 29-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 17-APR-2002; 2002US-0373260P.
PR 19-APR-2002; 2002US-0373815P.
PR 19-APR-2002; 2002US-0373817P.
PR 19-APR-2002; 2002US-0373884P.
PR 22-APR-2002; 2002US-0374977P.
PR 16-MAY-2002; 2002US-0381037P.
PR 16-MAY-2002; 2002US-0381038P.
PR 17-MAY-2002; 2002US-0381042P.
PR 28-MAY-2002; 2002US-0383656P.
PR 25-JUN-2002; 2002US-0391335P.
XX (SMIT//) SMITHSON G.
PA (MILL//) MILLET I.
PA (PEYM//) PEYMAN J. A.
PA (KEKU//) KEKUDA R.
PA (JUJJ//) JU J.
PA (LILL//) LI L.
PA (GUOX//) GUO X.
PA (PATT//) PATTURAJAN M.
PA (SPYT//) SPYTEK K. A.
PA (EDIN//) EDINGER S. R.
PA (ELLE//) ELLERMAN K.
PA (MALY//) MALYANKAR U. M.
PA (ORTT//) ORT T.
PA (GORM//) GORMAN L.
PA (ZERH//) ZERHUSEN B. D.
PA (ANDE//) ANDERSON D. W.
PA (ZHON//) ZHONG M.
PA (CATT//) CATTERTON E.
PA (JIWW//) JI W.
PA (MILL//) MILLER C. E.
PA (RAST//) RASTELLI L.
PA (STON//) STONE D. J.
PA (PENA//) PENA C. E. A.
PA (SHEN//) SHENOY S. G.
PA (SHIM//) SHIMKETS R. A.
PA (ROTH//) ROTHENBERG M. E.
PA (LEAC//) LEACH M. D.
PA (AGEE//) AGEE M. L.
PA (BERG//) BERGHS C.
PA (DIP1//) DIPIPPO V. A.
PA (EISE//) EISEN A.
PA (GANG//) GANGOLLI E. A.
PA (RIEG//) RIEGER D. K.
PA (SPAD//) SPADERNA S. K.
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI Shmkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;
XX
DR WPI; 2004-213931/20.
XX P-PSDB; ADN62845.
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,
XX diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.
PS Claim 20; SEQ ID NO 39; 395pp; English.
XX The invention relates to isolated NOVX polypeptides and polynucleotides.
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or
CC treat a medical condition in human related to the aberrant expression and
CC activity of NOVX polypeptides. For example, NOVX polypeptides and
CC polynucleotides may be used to treat disorders associated with decreased
CC expression or activity of NOVX by supplementing the patient our
CC production or to rectify mutations. Conversely, antisense NA molecules
CC may be administered to down regulate expression of NOVX polypeptides by
CC binding with the cells own genes and preventing their expression. NOVX
CC polynucleotides and complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar sequences in samples, and so which patients may be in need of
CC restorative therapy. NOVX polypeptides may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of NOVX. The
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be
CC used to modulate NOVX polynucleotide expression and activity of NOVX
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX
CC polypeptides and polynucleotides may be used in this way to prevent,
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
CC haematopoietic disorders, and the various dyslipidaemias, metabolic
CC disturbances associated with obesity, the metabolic syndrome X and
CC wasting disorders associated with chronic diseases and various cancers.
CC They may also be used as antibacterial agents. The present sequence
CC represents DNA encoding a human NOVX protein.
XX
SQ Sequence 1740 BP; 388 A; 477 C; 530 G; 345 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.18e-165 Length: 1740
Score: 1595.50 Matches: 312
Percent Similarity: 96.02% Conservative: 2
Best Local Similarity: 95.41% Mismatches: 1
Query Match: 92.92% Indels: 13
DB: Gaps: 1

US-10-631-958-2 (1-326) x ADN62844 (1-1740)
Qy 1 ProlYshHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyArgIle 20
Db 463 CCAAGCATTACTGGTATTATTCACCCGTTTCGAGGAGAAAGGACAAGCGGATA 522
Qy 21 TyrGluArgLysValAlaProLeuPheThrIleuAlaSerIleThrAspIleIleGly 40
Db 523 TATGAAAGAAAGTGGCACCACCTGTTCCACCTTACCTCCATCACCACCTGATCATC--- 579
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 580 -----GTTACTGAACATGCTATATCAGCCCAAGGAG 609
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp-Gl 80
Db 610 ACTCTGTATGAGATTACATAGACAATACACGCGCAT-GTCTGTGTCTGCGGAGATCGG 668
Qy 80 yMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAs 100
Db 669 TATGTTCAGCGAGGTGTCACGCTCTGATTGGGAGGACGCGAGAGGCGCGCGGGTGA 728
Qy 100 pGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAl 120
Db 729 CCAGAACACCCCGGGCTGTGCTGGTCCCGAGTAGCTCCGGATTGGATTCATCCCGC 788

```

QY 120 aGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAl 140
Db 789 AGGGTCAACGGACTCGTGTGTACTCCACCGTGGGCACCGACGCGAGAAACCTCGGC 848
QY 140 aLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisAsnSe 160
Db 849 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTAGTCCACACACAG 908
QY 160 rThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 180
Db 909 CACACTCCTCGTACTCCTCGTGTCCCTGGCTACGGCTTCTACGGGGACATCAAC 968
QY 180 sAspSerGluValAspArgTyrLeuAlaArgTyrAspPheSerGlyLeuIleYth 200
Db 969 GCACAGTGAAGAACAACGGTGTGGTCTCCAGATACGACTTTTCAGGTTTAAAGAC 1028
QY 200 rPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVa 220
Db 1029 CTTCTCTCCACCACTGCTATGAGGACAGTGTCTTCTCTCCCTGCACACACACGGT 1088
QY 220 lGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIy 240
Db 1089 GGGATCTCCAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGCGGCAAGCAA 1148
QY 240 sGlnGlnLeuGluGluGlnLysIleValLeuTyrGlyLeuGluAlaAlaGluaspVa 260
Db 1149 GCAGCAGCTGGAGGAGGAGCAAGAAGACATGTATGTTTGGAAAGCTGCGGAGCGT 1208
QY 260 lGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCy 280
Db 1209 GGAGAGTGGCAAGTCGTCTGTGGAGATTTCTGGCCATCAATGCCAACAATGTCCTG 1268
QY 280 sAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSe 300
Db 1269 TGCTTGTGCGCGAGGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 1328
QY 300 rAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHi 320
Db 1329 TGACCTCATCTCTCATCCGGAATGTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 1388
QY 320 sThrAsnGlnInAspGln 326
Db 1389 CACCAACCCAGCAGGACCAG 1407
RESULT 12
ID ADS10370
XX ADS10370 standard; DNA; 4702 BP.
XX AC ADS10370;
XX DT 16-DEC-2004 (first entry)
XX DE Human therapeutic DNA - SEQ ID 607.
XX KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
XX OS Homo sapiens.
XX PN WO2004080148-A2.
XX PD 23-SEP-2004.
XX PF 30-SEP-2003; 2003WO-US030720.
XX PR 02-OCT-2002; 2002US-0416186P.
XX PA (NUVE-) NUVELO INC.
XX PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX

DR WPI; 2004-668857/65.
DR P-PSDB; ADS11054.
XX New polynucleotide, useful in preparing a composition for diagnosing or
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT aplastic anemia or cancer for promoting wound healing.
XX Claim 1; SEQ ID NO 607; 718pp; English.
XX The invention relates to a novel isolated polynucleotide and the encoded
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during genetic therapy
CC procedures. The current sequence is that of a human therapeutic DNA of
CC the invention. The current sequence is not shown explicitly within the
CC specification but can be accessed from the WIPO web-site.
XX
SQ Sequence 4702 BP; 970 A; 1226 C; 1364 G; 1118 T; 0 U; 24 Other;
Alignment Scores:
Pred. No.: 4, 95e-145 Length: 4702
Score: 1414.50 Matches: 276
Percent Similarity: 80.58% Conservative: 2
Best Local Similarity: 80.00% Mismatches: 0
Query Match: 82.38% Indels: 67
DB: 13 Gaps: 1
US-10-631-958-2 (1-326) x ADS10370 (1-4702)
QY 49 GluValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAsp 68
Db 1627 AAGCTGATTACTGAACATGCTAATCAGSCCAAGGAGACTCTGTATGAGATTAAACATAGAC 1686
QY 69 LysTyrAspGlyIleValCysValGlyCysGlyMetPheSerGluValLeuHisGly 88
Db 1687 AAATACAGCGCATCGTCTGTGTGGCGAGATGATGTATGTTTCAGCGAGGTGTCGACGGT 1746
QY 89 LeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeu 108
Db 1747 CTGATTGGAGAGCAGCAGAGGAGCGCGGGTGCACACAGAACCCCGGCTGTGCTG 1806
QY 109 ValProSerSerLeuArgIleGlyIleIleProAla 120
Db 1807 GTCCCCAGTAGCTCCGGATTGGAATCATTTCCCGCAGGCTTTTGCAATGACCTGCGCAGGG 1866
QY 120 120
Db 1867 GACGAGGTGTCTGTCTCTCTGGCCCTGTGTCTGGCCCCGAGGGTGGCGCATGGTGCAC 1926
QY 120 120
Db 1927 ACTTTCATCTCGTTCACAGCTCTGGGATGTGAGCACCGCAGTCATCCCATTTTATGGAT 1986
QY 121 121
Db 1987 GAAGACAGGAGGACTGGGGACATGTGGCCCCCGGTGAGAACGCTGGTGGCTTGGACGGG 2046
QY 122 SerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeu 141
Db 2047 TCAACGGACTCGTGTGTGTACTCCACCGTGGCACCAGCGACGAGAAACCTCGCGCGCTG 2106
QY 142 HisIleValValGlyAspSerLeuAlaMetAspValSerValHisHisAsnSerThr 161
Db 2107 CATATCGTTGTGGGACTCGCTGGCCATGATGTCTCAGTCCATCCACCAACAGCACA 2166
QY 162 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAsp 181
Db 2167 CTCCTCTCGTACTCGTGTCTCTGCTGGCTACGGCTTCTACGGGGACATCATCAAGGAC 2226
QY 182 SerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 201

QY 208 GluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLys 227
 DB 1552 GAAGGACAGTGTCTCTCCCTCCACACACACCGTGGGATCTCCAAAGGAATAGGAAG 1611
 QY 228 ProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGln 247
 DB 1612 CCTCGCGGAGGATGCTTGTGTGAGGCAAGCAGCAGCTGGAGAGGAGCAG 1671
 QY 248 LysLysAlaLeuTyrGlyLeuGluAlaGluAspValGluGluTrpGlnValValCys 267
 DB 1672 AAGAAGCACTGTATGTTTGAAGCTGGAGGACGTGGAGAGTGGCAAGTCTCTGT 1731
 QY 268 GlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArg 287
 DB 1732 GGAAGATTCTGGCCATCAATGCCAACATGTCTGTGTGTGGCGGAGCCCCAGG 1791
 QY 288 GlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgLys 307
 DB 1792 GGCCTCTCCCGGCTGCCCACTTGGAGAGCGGTCTTCTGACCTCATCTCATCCGGAA 1851
 QY 308 CysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGln 326
 DB 1852 TGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGCACACCAACAGCAGGACCAG 1908
 RESULT 14
 AAS77730
 ID AAS77730 standard; cDNA; 2186 BP.
 XX
 AC AAS77730;
 DT 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #13534.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR P-PSDB; ABG13543.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1: SEQ ID NO 13534; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2186 BP; 455 A; 637 C; 669 G; 425 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.81e-135 Length: 2186
 Score: 1323.50 Matches: 275
 Percent Similarity: 58.26% Conservative: 0
 Best Local Similarity: 58.26% Mismatches: 0
 Query Match: 77.08% Indels: 197
 DB: 5 Gaps: 3

US-10-631-958-2 (1-326) x AAS77730 (1-2186)

QY 52 ThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAsp 71
 DB 3 ACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGAC 62
 QY 72 Gly----- 72
 DB 63 GGATGCTACATGCCCATTTCAAGGATGCCCTGTCTCGAAACCACAGCCCCGCTGGGAAG 122
 QY 72 ----- 72
 DB 123 GCTGCCCGGTGCAAGGTGTAGGCTACGGGGAAGGGCAAGAGCCTTCCCATGGAGCG 182
 QY 72 ----- 72
 DB 183 GGGCCAAAGCAGATGCCCTGCGCAAGGGCCAGGTGTGCAGGTCGCCCTCTCTGAAGCTGAA 242
 QY 72 ----- 72
 DB 243 GCCTCAGGCTGTGCGGCTCAGAGGGCGGCGACCTGCAGGAGCCCTGCACCGCTGTGTTC 302
 QY 73 -----lleValCysValGlyAspGlyMetPheSerGluValLeuHisGlyLeu 89
 DB 303 TCCTGCAGCATCGTCTGTGTCGGCGAGATGGTATGTTTCAGCGAGGTCTGCACGGTCTG 362
 QY 90 IleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuVal 109
 DB 363 ATGGGAGGACGACAGAGGAGCGCGGGTCCACAGAACCCACCCCGGGCTGTGTGTGTC 422
 QY 110 ProSerSerLeuArgIleGlyIleProAla----- 120
 DB 423 CCCAGTAGCCTCCGGATTGGAAATCATTTCCCGCAGGCTTTGCCAAATGACCTGGCAGGGGAC 482
 QY 120 ----- 120
 DB 483 GAGGTGTCTGTCTCTCTGCGCCCTGTGTCTGCGCCCGAGGGTGGCGGCATGTTGCACACT 542
 QY 120 ----- 120
 DB 543 TTCACCTCTCGTFCACAGCTCTGGGATGTGAGCACCGCAGTCATCCCCATTTTATGAGTAA 602
 QY 121 -----GlySer 122
 DB 603 GACAGGAGGACTGGGAGCATGTGGGCGCCCGGTGAGAACGCTGTGTGGCTTCGACGGGTCA 662
 QY 123 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 142
 DB 663 ACGGACTGGGTGTGTACTCCACCGTGGGACACGACGACGACGAGAACCTTCGCGCTGCAT 722
 QY 143 IleValVal----- 145

```
Db 723 ATCGTTGTTGCTGCTGCCCCAGGCTCGGAACACCCCGCGCTACTGCTGT 782
Qy 145 ----- 145
Db 783 GGCAGTGGGACGGGACAGCTGTGCTGGCTGGCCAGCTGTGGAACGATGCTCTGTGAG 842
Qy 146 -----GlyAspSerLeuAlaMetAspValSer 154
Db 843 GCCTCGAGGCTTCAGTCCAGGATCAGAGCCCCGGGACTCGCTGGCCATGGATGTCTCC 902
Qy 155 SerValHisHisAenSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 174
Db 903 TCAGTCCACCACAACAGCACACTCTTCGCTACTCCGTGTCCCTGCTGGCTACGGCTTC 962
Qy 175 TyrGlyAspIleIleValAspSerGluLysAspArgTrpLeuGlyLeuAlaAaTyrAsp 194
Db 963 TACGGGGACATCATCAAGGACAGTGAGAGAAACGGGTGGTCTTGCAGATACGAC 1022
Qy 195 PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu 214
Db 1023 TTTTCAGGTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTC 1082
Qy 215 ProLaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 234
Db 1083 CCTGCACAACACACGGTGGGATCTCCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTT 1142
Qy 235 ValCysArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrGlyLeu 254
Db 1143 GTTTGACGGCAAGCAGACAGCAGCTGGAGGAGAGCAGAAAGCACTGTATGGTTTG 1202
Qy 255 GluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsn 274
Db 1203 GAAGTCGGGAGGAGCTGGAGAGTGGCAAGTCTGTGGGAAGTTCTGGCCATCAAT 1262
Qy 275 AlaThrAenMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHis 294
Db 1263 GCCAACAACATGTCCTGTGCTGTGCGCGAGCCACAGGGCCCTCTCCCGGCTGCCAC 1322
Qy 295 LeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeu 314
Db 1323 TTGGGAGACGGGTCTTTCAGCTCTATCCTATCCGGAATGCTCCAGGTTCAATTTCTG 1382
Qy 315 ArgPheLeuIleArgHisThrAsnGlnAspGln 326
Db 1383 AGATTTCATCAGGCACACCAACAGCAGGACAG 1418

RESULT 15
AAC76031
ID AAC76031 standard; cDNA; 753 BP.
AC AAC76031;
XX
XX
DT 08-FEB-2001 (first entry)
XX
XX Human ORFX ORF1586 polynucleotide sequence SEQ ID NO:3171.
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antiporiatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antitviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
XX Homo sapiens.
```

```
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US008621.
XX
PR 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR P-PSDB; AAB41822.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 5; Page 2392; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antiporiatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antirheumatic; antithyroid; antianemic. The
CC antitviral; antifungal; antirheumatic; antithyroid; and antianemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 753 BP; 165 A; 195 C; 207 G; 186 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.72e-106 Length: 753
Score: 1055.00 Matches: 199
Percent Similarity: 99.50% Conservative: 0
Best Local Similarity: 99.50% Mismatches: 1
Query Match: 61.44% Indels: 0
DB: 3 Gaps: 0

US-10-631-958-2 (1-326) x AAC76031 (1-753)
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 133 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACCGACGACGAGAAACCTTCGGCG 192
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHisAsnSer 160
Db 193 CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAGC 252
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 253 ACACCTCTCGTACTCCGTGTCCCTGCTGGCTTCTACGGGAGCATCATCAAG 312
Qy 181 AspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 313 GACAGTGAGAAGAAACGGTGGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGACC 372
```

```
Qy 201 PheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db |||||
373 TTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACGGTG 432
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db |||||
433 GGATCTCCAAAGGGATAGGAGCCCTGCCGGGACAGGATGCTTTGTTTGACGCAAGCAAG 492
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db |||||
493 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAGGATGTG 552
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db |||||
553 GAGGAGTGCAAGTCGTCTGTGGGAAGTTTCTGGCCCATCAATGCCACAAACATGTCCTGT 612
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db |||||
613 GCTTGTGCGCGGAGCCCCAGGGGCTCTCCCCGGCTGCCCACTTGGGAGACGGGTCTTCT 672
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db |||||
673 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCTGGCAC 732
```

Search completed: September 6, 2005, 11:24:23
Job time : 542.298 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 18:18:34 ; Search time 5901.4 Seconds
(without alignments)
10410.363 Million cell updates/sec

Title: US-10-631-958-9
Perfect score: 1614
Sequence: 1 atggggggcagggggcgcc.....agccagactcacacagctga 1614

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hlc:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gse1:*
- 9: gb_gse2:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1116.8	69.2	1824	3	AK042077	Mus muscu
2	1115.2	69.1	4248	3	AK052269	Mus muscu
3	861.8	53.4	1063	4	BM479389	AGENCOURT
4	769.4	47.7	1059	5	BQ054406	AGENCOURT
5	744.4	46.1	1078	5	BQ063738	AGENCOURT
6	739.6	45.8	797	7	BQ000755	AGENCOURT
7	713	44.2	1047	5	BQ057191	AGENCOURT
8	704.4	43.6	713	7	CK000525	AGENCOURT
9	700	43.4	820	6	CD655311	AGENCOURT
10	684.4	42.4	732	7	CF135528	UI-HF-BNO
11	653	40.5	653	5	BX952302	DKF2p781l
12	649	40.2	661	7	CK296332	170005321
13	579.4	35.9	581	5	BP224560	BP224560
14	579.2	35.9	584	5	BP310011	BP310011
15	570.4	35.3	584	5	BP309990	BP309990
16	563.8	34.9	573	7	CF138275	UI-HF-BNO
17	551.2	34.2	758	6	CB246749	UI-M-F10-
18	543.2	33.7	550	7	CF138634	UI-HF-BNO
19	535.8	33.2	541	7	CK296331	170005321
20	504	31.2	582	5	BP274786	BP274786
21	491	30.4	812	7	CK603033	AGENCOURT
22	490.2	30.4	1017	5	BQ879739	AGENCOURT
23	477.2	29.6	488	5	BH430459	UI-HF-BNO
24	475	29.4	550	2	AW499858	UI-HF-BNO

25	464.4	28.8	474	2	AM503999	UI-HF-BNO
26	454	28.1	763	7	CK364327	AGENCOURT
27	442	27.4	449	2	AM503483	UI-HF-BNO
28	433.8	26.9	607	6	CB581157	AMGNNUC:N
29	429.2	26.6	653	6	BY754042	BY754042
30	428	26.5	688	7	CN296313	170005336
31	427.2	26.5	608	6	CB580936	AMGNNUC:N
32	404.2	25.0	662	2	BB638867	BB638867
33	402.8	25.0	576	6	CB608292	AMGNNUC:N
34	394	24.4	656	6	BY739924	BY739924
35	382.2	23.7	540	6	CB612897	AMGNNUC:N
36	381.4	23.6	553	6	CA578972	CA578972
37	376.6	23.3	545	6	CB611947	AMGNNUC:N
38	367.4	22.8	376	5	BH430866	UI-HF-BNO
39	366.4	22.7	680	2	BB235873	BB235873
40	359.2	22.3	385	5	BH428707	UI-HF-BNO
41	358.8	22.2	779	5	BH265377	603374309
42	353.6	21.9	819	7	CN235868	CN235868
43	349	21.6	759	5	BH264162	603814452
44	348.8	21.6	359	5	BH428926	UI-HF-BNO
45	333.2	20.6	582	4	BM735016	MONOL_23_

ALIGNMENTS

RESULT 1	AK042077	1824 bp	linear	HTC 03-APR-2004
LOCUS	AK042077	Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630056D11 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog (Homo sapiens), full insert sequence.		
ACCESSION	AK042077			
VERSION	AK042077.1	GI:26334912		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
AUTHORS	1	Normalizaton and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
TITLE	Genome Res. 10 (10), 1617-1630 (2000)			
JOURNAL	20499374			
MEDLINE	11042159			
PUBMED	11042159			
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
AUTHORS	1	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
TITLE	Genome Res. 10 (11), 1757-1771 (2000)			
JOURNAL	20530913			
MEDLINE	11076861			
PUBMED	11076861			
REFERENCE	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
AUTHORS	1	Functional annotation of a full-length mouse cDNA collection		
TITLE	Nature 409, 685-690 (2001)			
JOURNAL				
REFERENCE	5			

AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
REFERENCE	6 (bases 1 to 1824)		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saichou, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN) Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]		
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/WRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers		
FEATURES	1..1824 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM.DB:A630056D11" /db_xref="taxon:10090" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stages="3 days neonate" 85..1680 /note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTR Q9UGES, evidence: FASTA, 86.5%ID, 100%length, match=945) putative" /codon_start=1 /protein_id="BAC31157.1" /db_xref="GI:26334913" /translation="MGANGAAEPLHSLVWVKRRCAVLSLEPARALLRWRSPEPGPSA PGADARVLVSEIIAIEEKDDCEKHSRWHKMFPAFTVHRVKRVRRHWRKWAR TFWGADEOLCHLQTLRGLLESITSRPHLLVFINPFGKGQKRIYEKTVAPLFTL ASITTEIITTHANCAKELVEINTDSYDGVICVCGDGMFSEVLHGVIGRTQOSAGID PNHRAVLVPTLRIGIIPAGSTDCVSVTGTNDARTSALHIIIGSLAIDVSVHY HNTLRYSVSLIGFGYGDLLIDSEKKRWMLVRYDFSGLKTFLSHQYVEGTLSEFLA QHTVSPRDNKPCRGACVFCROSKQLEEEKALYLENAEBEWEQVTCGKFLAIN ATNMSPCPRPGSLSPFALGDGSDLLIRKCSRNFRLRIHRNOEQDFTFV EYVRVKFHFSTKHEVEDNDSDKEQKQKFGKIKDRPSCSTCSASRSSWNGDGVMSH PAIEVRVHCQLVRLFARGIEEES"		
CD5	1 ATGGGGGCGACGGGGCGGGCGGAGCCGCTGCAATCCGTGCTGGTGAAGACGACGCGC 60 85 ATGGGGGCAATGGGGCGGGCGGAGCCGCTGCAATCCGTGCTGGTGAAGACGACGCGC 144 61 TGCGCCGTGAGCCCTGGAGCCCGCGCGGGCTCTGCTGGCTGGTGGCGGAGCCCGGGGCC 120 145 TGTGCCCTGACGCTGGAGCCCGCGCGGGCTCTGCTACGCTGGTGGCGGAGCCCGGGAGC 204 121 GGAGCCGGCGCCCGCGCGGGAGTCCGTGCTGCTGTGCTCTATCTGAGATCATCGCCGTT 180 205 GGGCCCTCCCGCGCGGGTCCGATGCCCGCTCGGTACTGGTTCGGAGATCATCGCTGTT 264 181 GAGGAAAACAGACGCTTCACGGGAAACATCAAGGAGCAGTGGAAAAATGGCAGAAAAAG 240 265 GAGGAAAACAGCAGCTCGGAGAAACACGCCCTCCAGTGGCGGATGATGCAATGAAAAAT 324 241 CCTTACGCTTTACAGTTCACTGTGTAAAGAGACGACGCGGACCCGCTGGAAGTGGCG 300 325 CCGTTCCATTCACAGTCCACCGTGTGAAGCGAGTACGACATCACCGCTGGAAGTGGCA 384 301 CAGGTGACTTCTGCTGCCAGAGGAGCAGTGTGCTGCTGCTGGCTGCAGACCCCTCGG 360 385 CGGTGACCTTCTGGAGCGCGGACGAGCAGTGTGTCACTGTGGCTGCAGACCCCTCCGT 444 361 GAGATGCTGGAGAAAGCTGACGCTCCAGACCAAAGCATTTACTGGTATTTATCAACCCGTTT 420 445 GGGCTGCTGGAGAGCCCTGACTTCAAGACCGAGCAGCTTGTGGTATTTATCAACCCCTTC 504 421 GGAGAAAAGAGCAAGGCAAGCGGATATATGAAAGAAAAGTGGCAACACCTTACCTTCA 480 505 GGAGGAAAGGTGAGGCAAGCGCATCTATGAAAAAACAGTGGCGCTCTGTTTACCTTG 564 481 GCCTCCATCACCCTGACATCATCTGTTACTGAAATCTAATCAGGCAAGAGGAGCTGTG 540 565 GCTTCCATCATCTACGAGATCATCATTTACAGAGATGCCAACCAAGCAAGAGACTTTA 624 541 TATGAGATTAACTAGACAAATACGACGGCATCGTGTGTGTCGGCGGAGATGTTATGTTTC 600 625 TACGAGATCAACACAGACAGCTATGATGGCATCGTGTGCTGAGTGGGAGCGGATGTTTC 684 601 AGCAGGTGCTGCACGCTGTGATTTGGAGAGCAGCAGAGGAGCGCGGGGTGCAGCAAC 660 685 AGCAGGTGCTGCATGGGTGATTTGGAGGAGCAGCAGAGCGCTGTGATCGACCCCAAT 744 661 CACCCCGGCGTGTGCTGCCAGTAGCTCCCGATTCGGAATCGAATCATTCGCCGAGGTC 720 745 CACCCCGAGCGGTGCTGGTGGCCAGTACCCCTCAGGATCGGCATCATACCCGAGGGTCC 804 721 ACGGACTGCGTGTGTTACTCCACCGTGGGCAACAGCGACGACAGAAAACCTCGGCGCTGCAT 780 805 ACAGATTGTGTGTTACTCAACAGTGGGCACAAACGACGACAGAGACATCGGCTTTGCAC 864 781 ATCGTTGTTGGGACTCGCTGGCCATGGATGTGCTCAGTCCACCAACAGACGACACTC 840 865 ATCATTTTGGGACTCCTGCGCAATAGAGTGTCTCTGTGCACTACTACCAATACACGCTG 924 841 CTTCGCTACTCCGTGCTCCGTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT 900 925 CTGGGCTACTCGGTTTCTCTGCTGGGTACGGTTTCTACGGGGACTTAATCAAGGACAGT 984 901 GAGAAGAAAACGTTGGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTTAAAGACCTTCTC 960 985 GAAAAGAAAACGTTGGATGGGCTCTGTCGGTATGATTTCTCAGGTTTGAAGACCTTCTC 1044 961 TCCACCACTGCTATGAAGGAGCAGTGTCTTCTCCCTCGCAACACAGCGGTGGATCT 1020 1045 TCTCATGACTATGAAGGAGACACTGTCTTCTCCCTCCAGCAGACGACGCGGTGGATCT 1104 1021 CCAGGGATAGGAGCCCTCGCGGAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG 1080 1105 CCACGGGACAATAAACCCCTGCGGGCTGGGTGCTTCTGTGTGAGGAGCAAGCAACAG 1164		
ORIGIN	Query Match 69.2%; Score 1116.8; DB 3; Length 1824; Best Local Similarity 81.2%; Pred. No. 1.3e-279; Matches 1309; Conservative 0; Mismatches 302; Indels 1; Gaps 1;		

```
QY 1081 CTGGAGGAGCAGAGAAAGCACTGTATGTTTGGAGCTGCGGAGGACCTGGAGGAG 1140
|||||
Db 1165 CTGGAAGAGAGAGAGAAAGCCCTGTATGCTGGAGAACGCCGGAATGGAAGAG 1224
|||||
QY 1141 TGGCAAGTCGTCTGTGGGAAGTTTCTGGGCATCAATGCCACAAACATGCTCCTGTCTGT 1200
|||||
Db 1225 TGGCAAGTCACATGTGGGAAGTTCTGGGCATCAATGCCACAAACATGCTCCTGTCTGT 1284
|||||
QY 1201 CCGCGAGAGCCAGAGGCGCTCTCCCGGCTGCCACTTGGGAGAGCGGCTTCTCGACCTC 1260
|||||
Db 1285 CCGCGAGAGCCCTGGGCGCTGTCCCATTTGCCATCTCGGAGATGGGTCTTCTGACCTC 1344
|||||
QY 1261 ATCTCATCGGAATGCTCAGGTTCAATTTCTGAGATTCTCATCGGACACCAAC 1320
|||||
Db 1345 ATCTCATCGGAAGTCTCAGGTTCAATTTCTGAGATTCTCATCGGACACCAAC 1404
|||||
QY 1321 CAGCAGGACAGTTTGACCTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
|||||
Db 1405 CAGGAGGACAGTTTGACCTTCACTTTTGTGAAGTTTATCGAGTCAAGAAATTCACCTC 1464
|||||
QY 1381 ACCTCGAAGCACATGAGGATGAGGACGACGACCTCAAGAGGGGGGGAAGAGCGCTTT 1440
|||||
Db 1465 ACCTCGAAGCACGTTGGAACGAGGACGACATGACTCGAAGGAACAAGAGAGAGAGTTT 1524
|||||
QY 1441 GGGCACATTTGACAGGACGACCCCTCTGCTGCTGACCGTCTCAACAGCTCTCGAAC 1500
|||||
Db 1525 GGGAGAGATCGAAGGACAGACCCCTCTGACCTTGTCTCAGCTCCAGAGCTCTCGAAC 1584
|||||
QY 1501 TCGCAGGGGAGTCTCTGCACAGCCCTGCTCATCGAGTCCAGAGTCCACTGCCAGCTGGTT 1560
|||||
Db 1585 TCGATGCGGAGTATGTCACAGCCCGCCATGAGTTCAGGTCATCTGCCAGCTGGTG 1644
|||||
QY 1561 CGACTCTTTGCGAGGATTTGAAG-AGAAATCGAAGCGAGACTCACACAGC 1611
|||||
Db 1645 CGCCTCTTTGCTCGGGAATCGAGGAGAGTATGAAGCAAGAACCCCAAGC 1696
|||||

RESULT 2
AK052269
LOCUS Mus musculus 4248 bp mRNA linear HTC 03-APR-2004
DEFINITION library, clone:D330016D08 product:DA59H18.2 (NOVEL PROTEIN SIMILAR
TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS)
(FRAGMENT) homolog [Homo sapiens], full insert sequence.
ACCESSION AK052269
VERSION AK052269.1 GI:26342491
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
TITLE Meth. Enzymol. 303, 19-44 (1999)
JOURNAL 92779253
MEDLINE 10349636
PUBMED
REFERENCE
2
Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
AUTHORS Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
TITLE Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
```

```
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
JOURNAL sequencing pipeline with 384 multicapillary sequencer
MEDLINE Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 20530913
REFERENCE 11076861
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
JOURNAL of 60, 770 full-length cDNAs
REFERENCE
6
Nature 420, 563-573 (2002)
AUTHORS
7
(bases 1 to 4248)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
FEATURES
Location/Qualifiers
source
1..4248
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="PANTOM,DB:D330016D08"
/db_xref="taxon:10090"
/clone="D330016D08"
/tissue_type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
260..1855
/notes="unnamed protein product; DA59H18.2 (NOVEL PROTEIN
SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED)
PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPR|09UG65,
evidence: FASTY, 86.5%ID, 100*length, match=945)
putative"
/codon_start=1
/protein_id="BAC34908.1"
/db_xref="GI:26342492"
/translaton="MGANGAEPHLSVLVWRRRCVSLSEPARALLRWRSPEPGPSA
PGDARVLSVEIIAEEKDDCEKHSGRWKNMFPFVTHRVKRVHHRWQVAV
TFWSADQLQMLQTLRGLLESITSPKLLVFINFGKGQKQKRYEKTVAFLFTL
ASITETIITEHANOAKETIVNTDSYDGVCGVCGMFSEVLHVGIRGTQOAGID
PNHPRAVLVPTSLRIGIIPAGSTDVCYSTVGTNDARTSALHIIIGDSLADVSSVH
HTNLLRSVLLGFGFDLIDKSEKRWGLVRYDFSGLTKTFLSHQYEGTSLFLPA
QHTVSPRDNKPCRGACFVCRQSKQLEEBEKALYGLNAAEMEEQVTCGKFLAIN
ATNMSCAPRSPGGLSPFAHLGDSGLIILIRKCSRFNLFRLIRHTNQEDQDFTFV
```

EVYRVKPHFTSKHVEYEDNDNSKEQKQFKGI CKDRPSTCTCSASRSSWNCDEVMHS
PAIBVRVHCQLVRLFARGIEES"

ORIGIN		Query Match 69.1%; Score 1115.2; DB 3; Length 4248;	
		Best Local Similarity 81.1%; Pred. No. 4.2e-279;	
		Matches 1308; Conservative 0; Mismatches 303; Indels 1; Gaps 1;	
Qy	1	ATGGGGGCGAGCGGGGCGGAGCCGCTGCAATCCGTCGTGCTGGGTGAAGCAGCAGCGC	60
Db	260	ATGGGGGCAATGGGGGCGGCGAGCCGCTGCACCTCCGTCGCTGGGTGAACGCGGAGCG	319
Qy	61	TGCGCGGTGAGCTGGAGCCGCGGGGCTCTGCTGGCGTGGTGGCGAGCCCGGGGCC	120
Db	320	TGTGCGGTGAGCTGGAGCCGCGGGGCTCTGCTACGCTGCTGGGAGCCCGGAGGCC	379
Qy	121	GGAGCGGCGCCCGGGCGGCGATGCTGCTCTGCTGCTGCTGATCTGAGATCATCGCGTT	180
Db	380	GGGCGCTCGGGCGGGGTGCCGATGCCCGCTCGGTACTGCTGCTGGAGATCATCGCTGT	439
Qy	181	GAGGAAACAGAGCTTCAGGGGAAACATCAAGGCGAGTGGAAATGCGAGAAATGGAAG	240
Db	440	GAGGAAACAGAGCTTCAGGAAACACAGCCCTCCAGTGGCCGATGGCAATAATGGAAT	499
Qy	241	CTTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCGCACCGCTGGAGGTGGCG	300
Db	500	CCGTTGCAATTCACAGTCCACCGTGTGAAGCGAGTACGACATCACCGCTGGAAGTGGCA	559
Qy	301	CAGTGACTTTCTGGTGTGACAGAGAGAGCTGTGTCACTTGTGGCTGCGAGCCCTGGG	360
Db	560	CGGGTGACCTTTCTGGAGCGCGAGCAGAGCTGTGTCACTGTGGCTGCGAGCCCTCGT	619
Qy	361	GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGCTATTATCAACCCGTTT	420
Db	620	GGGCTGCTGGAGAGCTGACTTCAAGACCAAGACATTTGCTGGTATTATCAACCCCTTC	679
Qy	421	GGAGGAAAGACAAAGCAAGCGGATATATGAAAGAAAGTGGCACCACCTTTTACCTTA	480
Db	680	GGAGGAAAGTTCAGGGCAAGCGCATCTATGAAAGAAACAGTGGCGCTCTGTTTACCTTG	739
Qy	481	GCCTCCATCACACTGACATCATGTTACTGAAACATGCTAATCAGGCGCAAGGAGACTGTG	540
Db	740	GCCTCCATCACACTGAGAGATCATATTACAGAGCATGCCAACCAAGCCAGGAGACTTTA	799
Qy	541	TATGAGATTACATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGGTATGTTTC	600
Db	800	TACGAGATCAACACAGACAGCTATGATGGCATCGTGTGCTGAGTGGGAGCGGATGTTTC	859
Qy	601	AGCGAGGTGCTGCACGCTCTGATTGGGAGGACGAGAGGAGCGCGGGGTGCAGCAGAAC	660
Db	860	AGCGAGGTGCTGCATGGGTGATTGGGAGGACGAGCAGAGCGGCTGGTATCGACCCCAAT	919
Qy	661	CACCCCGGGCTGTGCTGGTCCCATGATAGCTTCGGGATTGGAATCATTTCCCGAGGGTCA	720
Db	920	CACCCCGAGCGGTGCTGGTCCCATGATAGCTTCGGGATTGGAATCATTTCCCGAGGGTCC	979
Qy	721	ACGAGCTCGGTGTACTTCCACCGTGGGACCGAGCGAGCAAACTCCGCGCTGCAT	780
Db	980	ACAGATTGTGTGTACTTCAACAGTGGGCAACACGACGAGAGACATCGGCTTTTGAC	1039
Qy	781	ATCGCTTTGTTGGGACTCGCTGCCATGATGTGCTCTCAGTCCACCAACAGCAGCATC	840
Db	1040	ATCATATTGGGACTCACTGGCAATAGACGTGCTCTGTGCACTACCAATAACACGCTG	1099
Qy	841	CTTGCTACTCGGTGCTCGTGGGCTTACGGCTTCTACGGGACATCATCAAGACAGT	900
Db	1100	CTGCGGTACTCGGTTTCTCTGCTGGGCTTACGGTTCCTACGGGAGCTTAATCAAGACAGT	1159
Qy	901	GAGAGAAACCGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTC	960
Db	1160	GAAAGAAACCGTGGATGGGCTCTGCGGTATGATTTTCTCAGGTTTGAAGACCTTCTC	1219
Qy	961	TCCACCACTGCTATGAAGGGACAGTGTCTTCTCTCTCTGCAACACACAGGTGGGATCT	1020

Db	1220	TCTCATCAGTACTATGAAGGACACGTCTCTCTCTCCAGCAGCAGCAGCGTGGGATCT	1279
Qy	1021	CCAGGGATAGGAGCCCTCCCGGAGGATGCTTTGTTGTCAGGCAAGCAAGCAGCAG	1080
Db	1280	CCAGGACAAATAAACCTCCCGGCTGGGTGCTTCTGTCGAGGAGAGCAAGCAACAG	1339
Qy	1081	CTGAGAGGAGCAGAGAAGAACACTGTATGGTTTGAAGCTCGGAGGACGTGGAGGAG	1140
Db	1340	CTGGAAGAAGAGGAAGAACCTGTATGGCTGGAGAACCCGAGGAATGGAAGAG	1399
Qy	1141	TGCAAGTCTGTGTGGAAAGTTTCTGGCCCATCAATGCCCAACAAATGTCTGTGCTGT	1200
Db	1400	TGGCAAGTGACATGTGGAAAGTTCTTGCCCATCAATGCCCAACAAATGTCTGTGCTGT	1459
Qy	1201	GGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTCTGACCTC	1260
Db	1460	CCTCGGAGCCCTGGGGCTGTCCCCATTTGCCCATCTGGGAGATGGGTCTTCTGACCTC	1519
Qy	1261	ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCACACCAAC	1320
Db	1520	ATCTTATCCGGAAGTCTCCAGGTTCACTTCTGAGATTCTCATCCGCGACACGAC	1579
Qy	1321	CAGCAGGACAGTTTGACTTCACTTTTGTGAAAGTTTATCGCTCAAGAAATTCAGTTT	1380
Db	1580	CAGGAGCAGGTTGAGCTTCACTTCTGTTGAAAGTTTATCGAGTCAAGAAATTCACCTT	1639
Qy	1381	ACGTCGAGCAGCATGGAGGATGAGGACGAGCCTCAAGGAGGGGGAGAGAGCGCTTT	1440
Db	1640	ACGTCGAGCAGCATGGGAAATACGAGGACAAATGACTCGAAGGAAACAAGAGGAGCAAGTTT	1699
Qy	1441	GGGCACATTTGCAGCAGCACCCCTCTCTGCTGTCACACCTCTCCACAGCTCTCTGGAAC	1500
Db	1700	GGGAGATCTGCAAGGACAGACCTCTTGCACCTTCTCAGCCTCCAGAGCTCTCTGGAAC	1759
Qy	1501	TGCGACCGGAGGTCTGCGACAGCCCTGCGCATCGAGGTGAGAGTCCATCTGCGAGCTGTT	1560
Db	1760	TGCGATCGGAGGTCTGCGACAGCCCGGCTTGGGTGAGGTGAGGTGAGGTGAGGTG	1819
Qy	1561	CGACTCTTTCAGCAGGAATTGAG-AGAAATCCGAGCGCAGACTCACACAGC	1611
Db	1820	CGCTCTTTTCTCGGGAAATCGAGGAAAGTATAGCAAGAAACCCCAAGC	1871
RESULT 3			
BM479389			
LOCUS			
DEFINITION			
AGENCOURT_6418742 NIH_MGC_67 Homo sapiens cdna clone IMAGE:550281			
5', mRNA sequence.			
ACCESSION			
BM479389			
VERSION			
BM479389.1 GI:18528431			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
REFERENCE			
1 (bases 1 to 1063)			
NIH-MGC http://mgi.nci.nih.gov/			
AUTHORS			
National Institutes of Health, Mammalian Gene Collection (MGC)			
TITLE			
Unpublished (1999)			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgabbs-remail.nih.gov			
Tissue Procurement: ATCC			
cdna Library Preparation: Life Technologies, Inc.			
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Agencourt Bioscience Corporation			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLAM12140 row: c column: 18			
High quality sequence stop: 665.			
Location/Qualifiers			
1. .1063			
FEATURES			
source			

/organism="Homo sapiens"									
/mol_type="mRNA"									
/db_xref="taxon:9606"									
/clone="IMAGE:5502281"									
/tissue_type="retinoblastoma"									
/lab_host="DH10B (phage-resistant)"									
/clone_lib="NIH_MGC_67"									
/notes="Organ: eye; Vector: pCMV-SPO6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.75 kb. Library constructed by Life Technologies."									
ORIGIN									
Query Match									
Best Local Similarity 53.4%; Score 861.8; DB 4; Length 1063;									
Matches 924; Conservative 0; Mismatches 42; Indels 10; Gaps 3;									
Qy	463	GCACCACTGTTACCTTAGCTCCATCACCACCTGACATCATCGTTACTGTAACATGCTAAT	522						
Db	1	GCACCACTGTTACCTTAGCTCCATCACCACCTGACATCATCGTTACTGTAACATGCTAAT	60						
Qy	523	CAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCATCTGCTGTGTC	582						
Db	61	CAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCATCTGCTGTGTC	120						
Qy	583	GGCGGAGATGGTATGTTTCAGCGAGGTCTGCGACGGTCTGATTTGGAGGAGCGCAGAGGAGC	642						
Db	121	GGCGGAGATGGTATGTTTCAGCGAGGTCTGCGACGGTCTGATTTGGAGGAGCGCAGAGGAGC	180						
Qy	643	GCCGGGGTGCACAGAACACCCCGGGCTGCTGCTGCCCTCAGTAGTACGCTCCGAGTTGGA	702						
Db	181	GCCGGGGTGCACAGAACACCCCGGGCTGCTGCTGCCCTCAGTAGTACGCTCCGAGTTGGA	240						
Qy	703	ATCATTTCCGCAAGGCTCAACGAGCTGCGTGTGTTACTTCCACCGTGGGCACCGACGACGCA	762						
Db	241	ATCATTTCCGCAAGGCTCAACGAGCTGCGTGTGTTACTTCCACCGTGGGCACCGACGACGCA	300						
Qy	763	GAACCTCGGGCTGCATATCGTTTGGGGACTCGCTGGCCATGGATGTCTCTAGTC	822						
Db	301	GAACCTCGGGCTGCATATCGTTTGGGGACTCGCTGGCCATGGATGTCTCTCTAGTC	360						
Qy	823	CACCACAAAGCAGCAGCTCTTCGCTACTTCGCTGTCCCTGCTGGGCTACGGCTCTACGGG	882						
Db	361	CACCACAAAGCAGCAGCTCTTCGCTACTTCGCTGTCCCTGCTGGGCTACGGCTCTACGGG	420						
Qy	883	GACATCATCAAGGACAGTGTGAGAGAAACGGTGTGGGTCTTGCCAGATACGACTTTTCA	942						
Db	421	GACATCATCAAGGACAGTGTGAGAGAAACGGTGTGGGTCTTGCCAGATACGACTTTTCA	480						
Qy	943	GGTTTAAAGACCTTCCTCTCCACCACTGCTATGAAGGGACAGTGTCTCTCTCCCTGCA	1002						
Db	481	GGTTTAAAGACCTTCCTCTCCACCACTGCTATGAAGGGACAGTGTCTCTCTCCCTGCA	540						
Qy	1003	CAACACACGGTGGGATCTCCAAGGGATAGGAAGCCCTGCCGGGCAGGATGCTTTGTTGC	1062						
Db	541	CAACACACGGTGGGATCTCCAAGGGATAGGAAGCCCTGCCGGGCAGGATGCTTTGTTGC	600						
Qy	1063	AGGCAAGCAGCAGCAGCTGGAGGAGCAGAGAAAGACATGTATGTTTGGAAAGCT	1122						
Db	601	AGGCAAGCAGCAGCAGCTGGAGGAGCAGAGAAAGACATGTATGTTTGGAAAGCT	660						
Qy	1123	GCGGAGACCTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGGCATCAATGCCACA	1182						
Db	661	GCGGAGACCTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGGCATCAATGCCACA	720						
Qy	1183	AACATGTCTGTGCTTCTGCGCGGAGGCCACAGGGGGCTCTCCCGGCTGCCACTTGGGA	1242						
Db	721	AACATGTCTGTGCTTCTGCGCGGAGGCCACAGGGGGCTCTCCCGGCTGCCACTTNGA	780						
Qy	1243	GACGGGTCTTCTGACCTCATCTCTCATCCGGAATATGCTCCAGGTTCAATTTCTGAGATT	1302						
Db	781	GACGGGTCTTCTGACCTCATCTCTCATCCGGAATATGCTCCAGGTTCAATTTCTGAGATT	840						

Qy	1303	CTCATCAGGCACACCAACCAAGCAGGACCAAGTTTGACTTCA--TTTGTGTTGAAGTTTATC 136
Db	841	CTCATCAGGCACACCAACCAAGCAGGACAGTTTGACTTCCCTTTGGTTGAAGTTTATC 900
Qy	1361	GCGTCAAG--AAATTCCAGTTTACGTCGAGCACATGAGGATG-----AGGACACGCA 1412
Db	901	CCCGTCAGGAAATCCAGTTTACCTCCAAACCCCTTGGAGGAGATGAAGGACACAGCA 960
Qy	1413	CCTCAAGGAGGGGGG 1428
Db	961	CCTCAAGGAGGGGGG 976
RESULT 4		
BQ054406		
LOCUS		
DEFINITION	BQ054406 1059 bp mRNA linear EST 29-MAR-2002	
ACCESSION	AGENCOURT_6771237 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5803668	
VERSION	BQ054406	
KEYWORDS	5', mRNA sequence.	
SOURCE	BQ054406.1 GI:19813746	
ORGANISM	Homo sapiens (human)	
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE	NIH-MGC http://mgc.nci.nih.gov/	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT	Unpublished (1999)	
	Contact: Robert Strausberg, Ph.D.	
	Email: cgapbs-r@mail.nih.gov	
	Tissue Procurement: Lou Staudt	
	cDNA Library Preparation: Rubin Laboratory	
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
	DNA Sequencing by: Agencourt Bioscience Corporation	
	Clone distribution: MGC clone distribution information can be	
	found through the I.M.A.G.E. Consortium/LLNL at:	
	http://image.llnl.gov	
	Plate: LLCM2040 row: a column: 13	
	High quality sequence stop: 601.	
FEATURES	Location/Qualifiers	
source	1..1059	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:5803668"	
	/tissue_types="lymphoma, cell line"	
	/lab_host="DH10B (phage-resistant)"	
	/clone_lib="NIH MGC 99"	
	/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:	
	EcoRI; cDNA made by oligo-dt priming, directionally cloned	
	into EcoRI/XhoI sites using the following 5' adaptor:	
	GGCAGCAG(G). Size-selected >500bp for average insert size	
	1.8kb. Library constructed by Ling Hong in the laboratory	
	of Gerald M. Rubin (University of California, Berkeley)	
	using ZAP-cDNA synthesis kit (Stratagene) and Superscript	
	II RT (Life Technologies). Note: this is a NIH_MGC	
	Library."	
ORIGIN		
Query Match	47.7%; Score 769.4; DB 5; Length 1059;	
Best Local Similarity	99.1%; Pred. No. 4.2e-189;	
Matches 784;	Conservative 0; Mismatches 6; Indels 1; Gaps 1;	
Qy	395	ATTTACTGGTATTTATCAACCCGTTTGGAGGAAAAGCAAGGACGGGATATATGAAA 454
Db	13	ATTTACTGGTATTTATCAACCCGTTTGGAGGAAAAGCAAGGACGGGATATATGAAA 72
Qy	455	GAAGATGGGCACACTGTTCCACCTTAGCTCCATCACCCTGCATCATCATCGTTACTGAAC 514
Db	73	GAAGATGGGCACACTGTTCCACCTTAGCTCCATCACCCTGCATCATCGTTACTGAAC 132
Ov	515	ATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGAGGCATCG 574

Db	133	ATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACAGCGCATCG	192
Qy	575	TCTGTGTCGGCGGAGATGTTATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGC	634
Db	193	TCTGTGTCGGCGGAGATGTTATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGC	252
Qy	635	AGAGGAGCGCCGGGGTTCAGCAGAACACACCCCGGGTGTCTGTGTCCTCCAGTAGCCCTCC	694
Db	253	AGAGGAGCGCCGGGGTTCAGCAGAACACACCCCGGGTGTCTGTGTCCTCCAGTAGCCCTCC	312
Qy	695	GGATTGGAATCATTTCCCGCAGGGTCAACGAGACTCGGTGTGTTATCTCCACCGTGGGACCA	754
Db	313	GGATTGGAATCATTTCCCGCAGGGTCAACGAGACTCGGTGTGTTATCTCCACCGTGGGACCA	372
Qy	755	GCGAGCAGAAACCTCGCGCTGCATATCGTTGTTGGGACTCGCTGCCATGATGTGT	814
Db	373	GCGAGCAGAAACCTCGCGCTGCATATCGTTGTTGGGACTCGCTGCCATGATGTGT	432
Qy	815	CCTCAGTCCACCAACAGACACACTCCTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCT	874
Db	433	CCTCAGTCCACCAACAGACACACTCCTTCGCTACTCGGTGTCCCTGCTGGGCTACGGCT	492
Qy	875	TCTACGGGACATCATCAAGGACAGTGAGAAAGAACCGTGTGGTCTTGCCAGATACG	934
Db	493	TCTACGGGACATCATCAAGGACAGTGAGAAAGAACCGTGTGGTCTTGCCAGATACG	552
Qy	935	ACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCC	994
Db	553	ACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCC	612
Qy	995	TCCCTGCACAAACACACCGTGGGATCTTCAAGGGATAGGAAGCCCTGCCGGGACGATGCT	1054
Db	613	TCCCTGCACAAACACACCGTGGGATCTTCAAGGGATAGGAAGCCCTGCCGGGACGATGCT	672
Qy	1055	TGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGAACACTGTATGGTT	1114
Db	673	TGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGAACACTGTATGGGT	732
Qy	1115	TGGAGCTGCGGAGGACGTGGAGGAGTGGCAAGTCTGTGGG-AGATTTCTGGCCATC	1173
Db	733	TGGAACTGCGGAGGACGTGGAGGAGTGGCAAGTCTGTGGGAAAGTTCTCTGGCCAT	792
Qy	1174	AATGCCACAAA 1184	
Db	793	AATGCCACAAA 803	
RESULT 5			
LOCUS	BQ063738		
DEFINITION	AGENCOURT_6873251 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925382		
	5', mRNA sequence.		
ACCESSION	BQ063738		
VERSION	BQ063738.1	GI:19891754	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	NIH-MGC http://mgi.nci.nih.gov/ .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: Lou Staudt		
	cDNA Library Preparation: Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		

FEATURES	source	Plate: LCM2096 row: p column: 23	
		High quality sequence stop: 640.	
	1. .1078	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5925382"		
	/tissue_type="lymphoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH_MGC_99"		
	Note: Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
ORIGIN			
	Query Match	46.1%; Score 744.4; DB 5; Length 1078;	
	Best Local Similarity	97.5%; Pred. No. 1.4e-182;	
	Matches 778; Conservative	0; Mismatches 16; Indels 4; Gaps 2;	
Qy	268	AAGAGACACAGCGGACCGCTGGAAGTGGCGCAGGTGACTTTCTGTGTCTCCAGAGGAG	327
Db	1	AAGAGACACAGCGGACCGCTGGAAGTGGCGCAGGTGACTTTCTGTGTCTCCAGAGGAG	60
Qy	328	CAGCTGTGTCACTTGTGGCTGCAGACCCCTCGGGAGATGCTGGAGAAAGTGCATCCAGA	387
Db	61	CAGCTGTGTCACTTGTGGCTGCAGACCCCTCGGGAGATGCTGGAGAAAGTGCATCCAGA	120
Qy	388	CCAAAGCATTTACTGTTTATCAACCCGTTTGGAGAAAAGACAAAGCGCGGATA	447
Db	121	CCAAAGCATTTACTGTTTATCAACCCGTTTGGAGAAAAGACAAAGCGCGGATA	180
Qy	448	TATGAAAGAAAGTGGCACCACCTGTTTCACTTACGCTCCATCACCCTGCATCATCGTT	507
Db	181	TATGAAAGAAAGTGGCACCACCTGTTTCACTTACGCTCCATCACCCTGCATCATCGTT	240
Qy	508	ACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGAC	567
Db	241	ACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGAC	300
Qy	568	GGCATCGTCTGTGTCGGCGGAGATGTTATGTTACGAGAGTGTCTGACGGTCTGATTGGG	627
Db	301	GGCATCGTCTGTGTCGGCGGAGATGTTATGTTACGAGAGTGTCTGACGGTCTGATTGGG	360
Qy	628	AGGACGACAGGAGCGCGGGGTGCACAGAACCCACCCCGGGGTGTGCTGGTCCCCAGT	687
Db	361	AGGACGACAGGAGCGCGGGGTGCACAGAACCCACCCCGGGGTGTGCTGGTCCCCAGT	420
Qy	688	AGCCTCCGGATTGGAATCATTTCCCGCAGGGTCAACGAGACTGCGTGTGTTACTCCACCGTG	747
Db	421	AGCCTCCGGATTGGAATCATTTCCCGCAGGGTCAACGAGACTGCGTGTGTTACTCCACCGTG	480
Qy	748	GGCACGACGACGACGAGAAACCTCGCGCTGCATATCGTTTGGGGACTCGCTGGCCATG	807
Db	481	GGCACGACGACGACGAGAAACCTCGCGCTGCATATCGTTTGGGGACTCGCTGGCCATG	540
Qy	808	GATGTGCTCTCAGTCCACCAACAGCAGACTCTCTCGTACTCCGTGTCTCCCTGCTGGGC	867
Db	541	GATGTGCTCTCAGTCCACCAACAGCAGACTCTCTCGTACTCCGTGTCTCCCTGCTGGGC	600
Qy	868	TACGGCTTCTACGGGACATCATCAAGGACAGTGCAGAGAAACCGGTGGTGGGTCTTGCC	927
Db	601	TACGGCTTCTACGGGACATCATCAAGGACAGTGCAGAGAAACCGGTGGTGGGTCTTGCC	660
Qy	928	AGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAGGACAGTGT	987
Db	661	AGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAGGACAGG	720

QY 988 TCCTTCTCTCCGACACACACAGG--TGGATCTCAAGGGA--TAGGAAGCCCTGCGG 1043
|||||
Db 721 TCCTTCTCTCCGACACACACAGGATCTCCAGGGAATAAGGAAGCCCTGCGC 780
|||||
QY 1044 GGCAGGATGCTTTGTTTG 1061
|||
Db 781 GGGCCAGGGATGCTTTTG 798
|||

RESULT 6
CK000755
LOCUS AGENCOURT_16363467 NIH_MGC_221 Homo sapiens cDNA clone
DEFINITION IMAGE:30707875 5', mRNA sequence.

ACCESSION CK000755
VERSION CK000755.1 GI:38526789
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 797)

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gethard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabbs-remail.nih.gov

Tissue Procurement: James Martin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: NDAM1073 row: c column: 20

High quality sequence stop: 656.

Location/Qualifiers

1..797

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30707875"

/lab_host="DH10B Tona"

/clone_lib="NIH MGC 221"

/note="Organ: mixed; Vector: pYX-Aac; Site_1: EcoRI;

Site_2: NotI; Library is oligo-dT primed and directionally

cloned. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated with

EcoR I adaptor, digested with Not I and then cloned

directionally into pYX-Aac vector. Average insert size

4-5kb. Adaptors 5' (AATTCGGCAGAGG)3' and 5'd

(CCTGTCGCG)3'. 3' Linker sequence - GCGGCGCTGAGGCC T18.

Sequencing primers 3'end: T3 promoter primer 5'd

(ATTAACCTCACTAAAGGGA)3'. 5' End: T7 promoter primer 5'd

(TAATACCACTCACTAAGGG)3'. Library was constructed in the

laboratory of M. Bento Soares. Note: this is a NIH_MGC

Library"

ORIGIN

Query Match 45.8%; Score 739.6; DB 7; Length 797;
Best Local Similarity 97.3%; Pred. No. 2.3e-181;
Matches 751; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 458 AAGTGGCACCACCTGTTACCTTAGCCTCCATCACCACCTGACATCATCGTTACTGAACATG 517
|||||
Db 13 ACGAGGCACCACCTGTTACCTTAGCCTCCATCACCACCTGACATCATCGTTACTGACATG 72
|||||

QY 518 CTAATCAGGCCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCT 577
|||||
Db 73 CTAATCAGGCCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCT 132
|||||
QY 578 GTGTCCGCGGAGATGTTATGTTTCAAGCAGAGGTGCTGCACGGTCTGATTGGGAGGACGAGA 637
|||||
Db 133 GTGTCCGCGGAGATGTTATGTTTCAAGCAGAGGTGCTGCACGGTCTGATTGGGAGGACGAGA 192
|||||
QY 638 GGAGCGCGGGGTGCAACAGAAACACCCCGGGCTGTGTCGGTCCCAAGTAGCTTCGGA 697
|||||
Db 193 GGAGCGCGGGGTGCAACAGAAACACCCCGGGCTGTGTCGGTCCCAAGTAGCTTCGGA 252
|||||
QY 698 TTGGAATCATTTCCCGCAGGGTCAACGGACTCCGTGTTGGGACTCGCTGGCCATGATGTGCT 757
|||||
Db 253 TTGGAATCATTTCCCGCAGGGTCAACGGACTCCGTGTTGGGACTCGCTGGCCATGATGTGCT 312
|||||
QY 758 ACGAGAAACCTCCGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTGCT 817
|||||
Db 313 ACGAGAAACCTCCGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTGCT 372
|||||
QY 818 CAGTCCACCAACAGCAGCACTCTTCTGCTACTCCGTGTCCTGTCCTGTCGGCTACGGCTTCT 877
|||||
Db 373 CAGTCCACCAACAGCAGCACTCTTCTGCTACTCCGTGTCCTGTCCTGTCGGCTACGGCTTCT 432
|||||
QY 878 ACGGGGACATCATCAAGGACAGTGAAGAACCGTGGTGGGCTTTGCCAGATACGACT 937
|||||
Db 433 ACGGGGACATCATCAAGGACAGTGAAGAACCGTGGTGGGCTTTGCCAGATACGACT 492
|||||
QY 938 TTTCAAGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCC 997
|||||
Db 493 TTTCAAGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCC 552
|||||
QY 998 CTGCACCAACACACGGTGGGATCTCCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTG 1057
|||||
Db 553 CTGCACCAACACACGGTGGGATCTCCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTG 612
|||||
QY 1058 TTTTCAGGCAAGCAAGCAGCAGTGGAGGAGGAGCAGAAAGCACTGTATGTTGG 1117
|||||
Db 613 TTTTCAGGCAAGCAAGCAGCAGTGGAGGAGGAGCAGAAAGCACTGTATGTTGG 672
|||||
QY 1118 AAGTTCGGAGGAGCTGGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGCCCATCAATG 1177
|||||
Db 673 AAGTTCGGAGGAGCTGGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGCCCATCAATG 732
|||||
QY 1178 CCACAAACATGTCCTGCTGTTGTCGCCGAGCCCGAGGGGCTCTTCCCGGC 1229
|||||
Db 733 GCACAAACATGTCCTGCTGTTGTCGCCGAGCCCGAGGGGCTCTCCCGCTGC 784
|||||

RESULT 7

BQ057191

LOCUS

DEFINITION

AGENCOURT_6769595 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5812382

5', mRNA sequence.

ACCESSION

BQ057191

VERSION

BQ057191.1 GI:19816531

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1047)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Ruben Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:


```

http://image.llnl.gov
Plate: LUCM2062 row: 1 column: 15
High quality sequence stop: 535.
Location/Qualifiers
1. .1047
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5812382"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 99"
/note="Organ: lymph; Vector: pOTB7; Site 1: xhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGACACGAG(G). Size-selected >500bp for average insert size
1.9kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

ORIGIN	Library	Query Match	Best Local Similarity	Score	DB 5;	Length	1047;
		Matches	825;	Conservative	0;	Mismatches	35;
					Indels	11;	Gaps
Qy	395	ATTTACTGGTATTTTATCAACCCGTTTGGAGGAAAGACAAAGCGGATATATGAAA	454				
Db	11	ATTTACTGGTATTTTATCAACCCGTTTGGAGGAAAGACAAAGCGGATATATGAAA	70				
Qy	455	GAANAAGTGGCACCACTGTTACCTTAGCTCCATCACCACTGACATCATCGTTACTGAAC	514				
Db	71	GAANAAGTGGCACCACTGTTACCTTAGCTCCATCACCACTGACATCATCGTTACTGAAC	130				
Qy	515	ATGCTAATCAGGCCAACGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCG	574				
Db	131	ATGCTAATCAGGCCAACGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCG	190				
Qy	575	TCTGTGTCGGCGGAGATGGTATGTTACGCGAGGTGTCGACGGTCTGATTGGGAGGACGC	634				
Db	191	TCTGTGTCGGCGGAGATGGTATGTTACGCGAGGTGTCGACGGTCTGATTGGGAGGACGC	250				
Qy	635	AGAGGAGCGCGGGTGCACAGAAACACCCCGGGCTGTCTGGTCCCGATGACCTCC	694				
Db	251	AGAGGAGCGCGGGTGCACAGAAACACCCCGGGCTGTCTGGTCCCGATGACCTCC	310				
Qy	695	GGATTGGAATCATTTCCGCGCAGGGTCAACGGACTCGGTGTGTTACTCCACCGTGGGCACCA	754				
Db	311	GGATTGGAATCATTTCCGCGCAGGGTCAACGGACTCGGTGTGTTACTCCACCGTGGGCACCA	370				
Qy	755	GGCAGCGACGAAACCTCGGCGCTGCATATCGTTGTTGGGGACTCCGCTGGCCATGGATGTGT	814				
Db	371	GGCAGCGACGAAACCTCGGCGCTGCATATCGTTGTTGGGGACTCCGCTGGCCATGGATGTGT	430				
Qy	815	CCTCAGTCCACCAACACAGCACCTCCTTCGTTACTCCGTTCCCTGCTGGGCTACGGCT	874				
Db	431	CCTCAGTCCACCAACACAGCACCTCCTTCGTTACTCCGTTCCCTGCTGGGCTACGGCT	490				
Qy	875	TCTACGGGGACATCATCAAGGACAGTGAAGAAACCGGTGGTGTGGGTCTTTGCCAGATACG	934				
Db	491	TCTACGGGGACATCATCAAGGACAGTGAAGAAACCGGTGGTGTGGGTCTTTGCCAGATACG	550				
Qy	935	ACTTTTCAGGTTTAAAGACCTTCTCTCCCACTGCTATGAA - GGGACAGTGTCTCTTC	993				
Db	551	ACTTTTCAGGTTTAAAGACCTTCTCTCCCACTGCTATGAAAGGGGACAGTGTCTCTTC	610				
Qy	994	CTCCCTGCACAAACACAGCGTGGGATCT - CCAAGGGATAGGAAGCCCTG - CCGGGCAGAT	1051				
Db	611	CTCCCTGCACAAACACAGCGTGGGATCTCCCAAGGGATAGGAACCTTCTCCCGGGCAGAT	670				
Qy	1052	GCTTTTG - TTTGACGGCAAGCAAGCAGCTGGAGGAGGACGAGAAAGC - ACTGT	1108				

Db	671	CTTTGTTTGGCGGAAACGACACGACCTGGAGGAGGAGCCCAAGAAACCACTTGT	730
Qy	1109	ATGCTTTGGAGCTGCGGAGGACGTTGGAGGAGTGGCAAGTCGTCT-GTGGGAAGTTTCTG	1167
Db	731	ATGCTTTGGAACTCTGCGGAAGACGCGGAGGAGTGGCAAACTCTCTGCGGGGGAAGTTTCTG	790
Qy	1168	GCCATCAATGCCCAACATGTCTCT-GTGCTTGTGCGCGGAGCCCGAGGGGCTCT--C	1223
Db	791	GCAATCATCGCCCAACACCTGTCTCTGTGTGTCTGTGCGCGGAACCCCGAGGGGCTTCTCC	850
Qy	1224	CCCGGCTGCGCCACTCTGGGAGACGGGTCTTCT	1254
Db	851	CCCGGCTGCCCAATTTGGGAGAAGGGTCTCT	881

RESULT 8

CK000525

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

713 bp mRNA linear EST 26-NOV-2003

AGENCOURT 16369000 NIH MGC 221 Homo sapiens cDNA clone

IMAGE:30708597 5', mRNA sequence.

CK000525

CK000525.1 GI:38526559

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 713)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapns@email.nih.gov

Tissue Procurement: James Martin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: NDAMI075 row: a column: 22

High quality sequence stop: 689.

Location/Qualifiers

1. 713

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30708597"

/lab_host="DH10B Tona"

/clone_lib="NIH_MGC_221"

/note="Organ: mixed; Vector: pYX-Asc; Site: 1: EcoRI; Site 2: NotI; Library is oligo-dT primed and directionally cloned Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. Average insert size 4-5kb. Adaptors 5' (AATTCGACAGG) 3' and 5' (CTCTGCGCG) 3'. 3' Linker sequence - GCGGCGCTGAGAGCC T18.

Sequencing primers 3' end: T3 promoter primer 5' (ATATACCTCTCAAGGGA) 3'. 5' End: T7 promoter primer 5' (TTTACGACTCTATAGG) 3'. Library was constructed in the laboratory of M. Bento Soares. Note: this is a NIH MGC Library"

FEATURES

source

ORIGIN	Library
Query Match	43.6%; Score 704.4; DB 7; Length 713;
Best Local Similarity	99.9%; Pred. No. 3.4e-172;
Matches 705; Conservative	0; Mismatches 1; Indels 0; Gaps 0;


```
QY 828 CAACAGCACACTCTTCCGCTACTCGTGTCCCTGCTGGCTACGGCTTCTACGGGGACAT 887
DB 8 CAACAGCACACTCTTCCGCTACTCGTGTCCCTGCTGGCTACGGCTTCTACGGGGACAT 67
QY 888 CATCAAGGACAGTGAAGAAACGGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTT 947
DB 68 CATCAAGGACAGTGAAGAAACGGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTT 127
QY 948 AAGACCTTCTCTCCACCACTGTATGAAGGGACAGTGTCTCTCCCTCGCACAACA 1007
DB 128 AAGACCTTCTCTCCACCACTGTATGAAGGGACAGTGTCTCTCCCTCGCACAACA 187
QY 1008 CACGCTGGGATCTCCAAGGGATAGAACCCCTGCGGGGAGAGATCTTTGTTGAGGCA 1067
DB 188 CACGCTGGGATCTCCAAGGGATAGAACCCCTGCGGGGAGAGATCTTTGTTGAGGCA 247
QY 1068 AAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCAGTGTATGTTTGGAGCTGCGGA 1127
DB 248 AAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCAGTGTATGTTTGGAGCTGCGGA 307
QY 1128 GGAACCTGGAGGAGTGGCAAGTCTGTGGGAAAGTTTCTGGCCATCAATGCCACAACAT 1187
DB 308 GGAACCTGGAGGAGTGGCAAGTCTGTGGGAAAGTTTCTGGCCATCAATGCCACAACAT 367
QY 1188 GTCTGTGCTGTGTCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGCGG 1247
DB 368 GTCTGTGCTGTGTCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGCGG 427
QY 1248 GTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTTCAATTTCTGAGATTTCTCAT 1307
DB 428 GTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTTCAATTTCTGAGATTTCTCAT 487
QY 1308 CAGGCACACCAACAGCAGGACCACTTTGACTTCACTTTTGTGAAGTTTATCGGTCAA 1367
DB 488 CAGGCACACCAACAGCAGGACCACTTTGACTTCACTTTTGTGAAGTTTATCGGTCAA 547
QY 1368 GAAATTCAGTTTACGTGAAGCAGATGAGGATGAGGACGACGACCTCAAGGAGGGGG 1427
DB 548 GAAATTCAGTTTACGTGAAGCAGATGAGGATGAGGACGACGACCTCAAGGAGGGGG 607
QY 1428 GAAGAAGCGCTTTGGGACATTTGAGCAGCAGCCACCCCTCTGCTGCTGCACCGTCTCAA 1487
DB 608 GAAGAAGCGCTTTGGGACATTTGAGCAGCAGCCACCCCTCTGCTGCTGCACCGTCTCAA 667
QY 1488 CAGCTCTCGAAGTCCGAGGGGGAGTCTTGACAGCCCTGCAATC 1533
DB 668 CAGCTCTCGAAGTCCGAGGGGGAGTCTTGACAGCCCTGCAATC 713
```

RESULT 9

```
CD655311
LOCUS CD655311 820 bp mRNA linear EST 18-JUN-2003
DEFINITION AGENCOURT_14552675 NIA Human H1 Embryonic Stem Cell cDNA Library
(Long) Homo sapiens cDNA clone IMAGE:30426593 5', mRNA sequence.
ACCESSION CD655311
VERSION CD655311.1 GI:31895467
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 820)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contract: Daniela S. Gerhard, Ph.D.

AUTHORS

Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov

JOURNAL

Tissue Procurement: Irene Ginis and Mahendra Rao, NIA

COMMENT

cDNA Library Preparation: Yulan Piao and Minoru Ko

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: NDAMS12 row: k column: 18

High quality sequence stop: 673.

FEATURES
source

Location/Qualifiers
1..820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30426593"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI;
This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01/H1 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID: 11544199] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lona-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

```
Query Match 43.4%; Score 700; DB 6; Length 820;
Best Local Similarity 97.1%; Pred. No. 4.9e-171;
Matches 712; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 738 CTCACCGTGGGACACGAGCGAGAAACCTCGGCGCTGCATATCGTTGTGGGACTC 797
DB 11 CTCACCGTGGGACACGAGCGAGAAACCTCGGCGCTGCATATCGTTGTGGGACTC 70
QY 798 GCTGGCCATGATGTGCTTCTACGGGAGCATCATCAAGGACAGTGCAGAGAAACCGTGGTT 917
DB 71 GCTGGCCATGATGTGCTTCTACGGGAGCATCATCAAGGACAGTGCAGAGAAACCGTGGTT 130
QY 858 CCTGCTGGGCTACGGCTTCTACGGGAGCATCATCAAGGACAGTGCAGAGAAACCGTGGTT 917
DB 131 CCTGCTGGGCTACGGCTTCTACGGGAGCATCATCAAGGACAGTGCAGAGAAACCGTGGTT 190
QY 918 GGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGA 977
DB 191 GGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGA 250
QY 978 AGGACAGTGTCTTCTCTCTGACACACACCGTGGGATCTCCAGGGATAGGAAGCC 1037
```

Db 251 AGGACAGTGTCTTCTCTCCCTGCACACACACCGTGGGATCTCAAAGGATAGGAAGCC 310
Qy 1038 CTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAA 1097
Db 311 CTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAA 370
Qy 1098 GAAAGCACTGATGTTTGGAGCTGCGGAGGACGTGGAGGAGTGGCAAGTCGTCTGTGG 1157
Db 371 GAAAGCACTGATGTTTGGAGCTGCGGAGGACGTGGAGGAGTGGCAAGTCGTCTGTGG 430
Qy 1158 GAAAGTTCTGCCATCAATGACCAAAATGCTCTGTGCTTGTGCGCGAGGCCCGAGGG 1217
Db 431 GAAAGTTCTGCCATCAATGACCAAAATGCTCTGTGCTTGTGCGCGAGGCCCGAGGG 490
Qy 1218 CCTCTCCCGGCTGCCACTTGGAGAGCGGTCTTCTGACCTCATCTCATCCGGAATG 1277
Db 491 CCTCTCCCGGCTGCCACTTGGAGAGCGGTCTTCTGACCTCATCTCATCCGGAATG 550
Qy 1278 CTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCAAGTTGA 1337
Db 551 CTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCAAGTTGA 610
Qy 1338 CTTACATTTTGTGAAGTTATCGGTCAAGAAATTCAGTTTACGTGCAAGCAGATGGA 1397
Db 611 CTTACATTTTGTGAAGTTATCGGTCAAGAAATTCAGTTTACGTGCAAGCAGATGGA 670
Qy 1398 GGATGAGACAGCAGCTCAAGAGGGGGGGAAGAGCGCTTGGGCACATTGCGACAGCAG 1457
Db 671 GGATGAGACAGCAGCTCAAGAGGGGGGGAAGAGCGCTTGGGCACATTGCGACAGCAG 730
Qy 1458 CCACCCCTCTGC 1470
Db 731 CTTCTGCTGCTGC 743

RESULT 10
CF135528
LOCUS
DEFINITION
IMAGE:3090211 5', mRNA sequence.
ACCESSION
CF135528.1 GI:33250972
VERSION
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (Bases 1 to 732)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
7044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Louis Staudt
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfi.html
Seg primer: pYX-5.
Location/Qualifiers
1..732
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:3090211"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LTI)"
/clone_lib="NIH_MGC_50"
/notes="Vector: p7713-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN

Query Match 42.4%; Score 684.4; DB 7; Length 732;
Best Local Similarity 99.7%; Pred. No. 5.5e-167;
Matches 696; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 918 GGGTCTTGGCAGATACGACTTTTAAAGACCTTCTCTCCACCACTGCTATGA 977
Db 8 GGGTCTTGGCAGATACGACTTTTAAAGACCTTCTCTCCACCACTGCTATGA 67
Qy 978 AGGACAGTGTCTTCTCCCTGCACACACCGTGGGATCTCAAAGGATAGGAAGCC 1037
Db 68 AGGACAGTGTCTTCTCCCTGCACACACCGTGGGATCTCAAAGGATAGGAAGCC 127
Qy 1038 CTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAA 1097
Db 128 CTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAA 187
Qy 1098 GAAAGCACTGATGTTTGGAGCTGCGGAGGAGTGGCAAGTCGTCTGTGG 1157
Db 188 GAAAGCACTGATGTTTGGAGCTGCGGAGGAGTGGCAAGTCGTCTGTGG 247
Qy 1158 GAAAGTTCTGCCATCAATGCCAACAATGCTCTGTGCTTGTGCGCGAGGCCCGAGGG 1217
Db 248 GAAAGTTCTGCCATCAATGCCAACAATGCTCTGTGCTTGTGCGCGAGGCCCGAGGG 307
Qy 1218 CCTCTCCCGGCTGCCACTTGGAGAGCGGTCTTCTGACCTCATCTCATCCGGAATG 1277
Db 308 CCTCTCCCGGCTGCCACTTGGAGAGCGGTCTTCTGACCTCATCTCATCCGGAATG 367
Qy 1278 CTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCAAGTTGA 1337
Db 368 CTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCAAGTTGA 427
Qy 1338 CTTACATTTTGTGAAGTTTATCGGTCAAGAAATTCAGTTTACGTGCAAGCAGATGGA 1397
Db 428 CTTACATTTTGTGAAGTTTATCGGTCAAGAAATTCAGTTTACGTGCAAGCAGATGGA 487
Qy 1398 GGATGAGACAGCAGCAGCTCAA- GGAGGGGGGGAAGAGCGCTTGGGCACATTGCGAGCA 1456
Db 488 GGATGAGACAGCAGCAGCTCAAAGGGAGGGGGGGAAGAGCGCTTGGGCACATTGCGAGCA 547
Qy 1457 GCCACCCCTCTGCTGTGTCACCGTCTTCCAAAGCTCTTGGAACTGCGAGGGAGGTCC 1516
Db 548 GCCACCCCTCTGCTGTGTCACCGTCTTCCAAAGCTCTTGGAACTGCGAGGGAGGTCC 607
Qy 1517 TGACACGCTTGCATCGAGTCTCAGCTCCACAGCTGTTTCGACTCTTTTCACAGAG 1576
Db 608 TGACACGCTTGCATCGAGTCTCAGCTCCACAGCTGTTTCGACTCTTTTCACAGAG 667
Qy 1577 GAATTGAAGAGATCCCAAGCCAGACTTCACACAGCTGA 1614
Db 668 GAATTGAAGAGATCCCAAGCCAGACTTCACACAGCTGA 705

RESULT 11
BX952302
LOCUS
DEFINITION
DXFZp781l11183 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
ACCESSION
BX952302
VERSION
BX952302.1 GI:43428907

```
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 653)
AUTHORS Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Pobo, G., Han, M. and Wiemann, S.
TITLE Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp781L1183) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..653
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781L1183"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/notes="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"
ORIGIN
Query Match 40.5%; Score 653; DB 5; Length 653;
Best Local Similarity 100.0%; Pred. No. 8.3e-159; Indels 0; Gaps 0;
Matches 653; Conservative 0; Mismatches 0;
QY 888 CATCAAGACAGCTGAGAGAAACGGTGGTCTTGCCAGATACGACTTTTCAGGTTT 947
DB 1 CATCAAGACAGCTGAGAGAAACGGTGGTCTTGCCAGATACGACTTTTCAGGTTT 60
QY 948 AAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTGCACAACA 1007
DB 61 AAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTGCACAACA 120
QY 1008 CAGCGTGGGATCTCCAGGATAGGAGCCCTGCCGGCAGGATGCTTTGTTTGAGGCA 1067
DB 121 CAGCGTGGGATCTCCAGGATAGGAGCCCTGCCGGCAGGATGCTTTGTTTGAGGCA 180
QY 1068 AAGCAAGACAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGTTTGGAGCTGCGGA 1127
DB 181 AAGCAAGACAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGTTTGGAGCTGCGGA 240
QY 1128 GCACGTGGAGAGTGGCAAGTCGTCTGTGGGAAGTTTTCGGCCATCAATGCCACAAACAT 1187
DB 241 GCACGTGGAGAGTGGCAAGTCGTCTGTGGGAAGTTTTCGGCCATCAATGCCACAAACAT 300
QY 1188 GTCTGTGTCTTTCGCGGAGGCCAGGGCCCTCTCCCGGCTGCCACTTGGGAGCGG 1247
DB 301 GTCTGTGTCTTTCGCGGAGGCCAGGGCCCTCTCCCGGCTGCCACTTGGGAGCGG 360
QY 1248 GTCTTCTGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTCTTGAGATTTCTCAT 1307
DB 361 GTCTTCTGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTCTTGAGATTTCTCAT 420
QY 1308 CAGGCACACCAACACGAGGACAGTTTGAATTTGATTTGTTGAGTTTATCGCGTCAA 1367
DB 421 CAGGCACACCAACACGAGGACAGTTTGAATTTGATTTGTTGAGTTTATCGCGTCAA 480
QY 1368 GAAATTCAGTTTACGTCTGAAGCACATGGAGATAGGACAGCGACCTCAAGGAGGGGG 1427
DB 480 GAAATTCAGTTTACGTCTGAAGCACATGGAGATAGGACAGCGACCTCAAGGAGGGGG 1427

Db 481 GAAATTCAGTTTACGTCTGAAGCACATGGAGATAGGACAGCGACCTCAAGGAGGGGG 540
QY 1428 GAAGAAGCGCTTTGGGCACATTTGCAGACGCCCTCTCTGCTGCTGCACCGCTCCAA 1487
DB 541 GAAGAAGCGCTTTGGGCACATTTGCAGACGCCCTCTCTGCTGCTGCACCGCTCCAA 600
QY 1488 CAGCTCTCTGGAACCTGCACCGGGAGGCTCTGCACAGCCCTGCCATCGAGGTCA 1540
DB 601 CAGCTCTCTGGAACCTGCACCGGGAGGCTCTGCACAGCCCTGCCATCGAGGTCA 653

RESULT 12
LOCUS CN296312 661 bp mRNA linear EST 16-MAY-2004
DEFINITION 17005332197874 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN296312
VERSION CN296312.1 GI:47312726
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 661)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
JOURNAL Contact: Brandenberger R
COMMENT Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 661 Std Error: 0.00.
FEATURES
Location/Qualifiers
1..661
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/notes="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from HES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."
ORIGIN
Query Match 40.2%; Score 649; DB 7; Length 661;
Best Local Similarity 99.8%; Pred. No. 9.2e-158;
Matches 660; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 830 ACAGACACTCTTCGCTACTCCGCTGCTTCCGCTGCTGGGCTACGGCTTACGGGACATCA 889
DB 1 ACAGACACTCTTCGCTACTCCGCTGCTTCCGCTGCTGGGCTTACGGGACATCA 60
QY 890 TCAAGGACAGTGAGAGAAACGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAA 949
DB 61 TCAAGGACAGTGAGAGAAACGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAA 120
QY 950 AGACTTCTCTCCACCACCTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAACA 1009
DB 121 AGACTTCTCTCCACCACCTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAACA 180
QY 1010 CGGTGGGATCTCCAAAGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTGAGGCAAA 1069
DB 181 CGGTGGGATCTCCAAAGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTGAGGCAAA 240
QY 1070 GCAAGCAGCTCGGAGGAGGAGAGAGACATGCTATGTTGTTGAACTGCGAGG 1129
DB 1129 GCAAGCAGCTCGGAGGAGGAGAGAGACATGCTATGTTGTTGAACTGCGAGG 1129
```

```
Db 241 GCAAGCAGCAGCTGGAGGAGGAGCAGAGAAGCACTGTATGTTTGGAAAGCTGCGGAGG 300
Qy 1130 ACCTGGAGGAGTGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCACAACATGT 1189
Db 301 ACGTGGAGGAGTGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCACAACATGT 360
Qy 1190 CCTGTGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGGCTGCCACTTGGGAGACGGGT 1249
Db 361 CCTGTGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGGCTGCCACTTGGGAGACGGGT 420
Qy 1250 CTTCAGACCTCATCTCATCCGGAATGCTCCAGGTTCATATTTCTGAGATTTCTCATCA 1309
Db 421 CTTCAGACCTCATCTCATCCGGAATGCTCCAGGTTCATATTTCTGAGATTTCTCATCA 480
Qy 1310 GGCACACCAACAGCAGGACCAAGTTGACTTTTGTGTTGAGTTTATCGCGTCAAGA 1369
Db 481 GGCACACCAACAGCAGGACCAAGTTGACTTTTGTGTTGAGTTTATCGCGTCAAGA 540
Qy 1370 AATTCAGTTTACGTTCGAAGCACATGGAGGATGAGGACGAGCCTCAAGGA-GGGGGGG 1428
Db 541 AATTCAGTTTACGTTCGAAGCACATGGAGGATGAGGACGAGCCTCAAGGAGGGGGGG 600
Qy 1429 AAGAAGCGCTTTGGGACATTTGAGCAGCAGCACCCCTCTCTGCTGTCGACCGTCTCAAC 1488
Db 601 AAGAAGCGCTTTGGGACATTTGAGCAGCAGCACCCCTCTCTGCTGTCGACCGTCTCAAC 660
Qy 1489 A 1489
Db 661 A 661

RESULT 13
BP224560
LOCUS BP224560 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA
DEFINITION clone DAT01067, mRNA sequence.
ACCESSION BP224560
VERSION BP224560.1 GI:52097465
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 581)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1 .581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DAT01067"
/cell_type="lymphocyte"
/cell_line="Daudi"
/clone_lib="Sugano cDNA library, lymphocyte Daudi"
/note="Burkitt's lymphoma"

ORIGIN
Query Match 35.9%; Score 579.4; DB 5; Length 581;
Best Local Similarity 99.8%; Pred. No. 1.2e-139;
Matches 580; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 548 TTAACATAGACAAATACACGGCATCGTCTGTGCGCGGAGATGGTATGTTACGCGAGG 607
Db 1 TTAACATAGACAAATACACGGCATCGTCTGTGCGCGGAGATGGTATGTTACGCGAGG 60
```

```
Qy 608 TGCTGACCGTCTGATTGGGAGGACGCGAGGAGCGCGGGGTGCGACCAAGAACCAACCCCC 667
Db 61 TGCTGACCGTCTGATTGGGAGGACGCGAGGAGCGCGGGGTGCGACCAAGAACCAACCCCC 120
Qy 668 GGGCTGTGCTGGTCCCAGTAGCCTCCGGATTGGAAATCATTTCCGCGAGGGTCAACGAGCT 727
Db 121 GGGCTGTGCTGGTCCCAGTAGCCTCCGGATTGGAAATCATTTCCGCGAGGGTCAACGAGCT 180
Qy 728 GCGTGTGTTACTCACCGTGGGACACGACGCGAGAAACCTCGGCGCTGCATATCGTTG 787
Db 181 GCGTGTGTTACTCACCGTGGGACACGACGCGAGAAACCTCGGCGCTGCATATCGTTG 240
Qy 788 TTGGGGACTCGCTGGCCATGATGTCTTCAGTCCACCAACAGCAGCAGCTCTCTCGCT 847
Db 241 TTGGGGACTCGCTGGCCATGATGTCTTCAGTCCACCAACAGCAGCAGCTCTCTCGCT 300
Qy 848 ACTCCGTGTCCTGCTGGGCTACGGCTTTACGGGGACATCATCAAGGACAGTGAGAAGA 907
Db 301 ACTCCGTGTCCTGCTGGGCTACGGCTTTACGGGGACATCATCAAGGACAGTGAGAAGA 360
Qy 908 AACGTTGGTTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCACC 967
Db 361 AACGTTGGTTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCACC 420
Qy 968 ACTGCTATGAAGGACAGTGTCTTCTCTCTGACAAACACACGCTGGGATCTCCAAAGG 1027
Db 421 ACTGCTATGAAGGACAGTGTCTTCTCTCTGACAAACACACGCTGGGATCTCCAAAGG 480
Qy 1028 ATAGGAAGCCCTCGCGGAGGATGCTTTGTCAGGCAAGCAAGCAGCAGCTGGAGG 1087
Db 481 ATAGGAAGCCCTCGCGGAGGATGCTTTGTCAGGCAAGCAAGCAGCAGCTGGAGG 540
Qy 1088 AGGAGCAGAGAAAGCACTGTATGTTTGAAGCTGCGGAG 1128
Db 541 AGGAGCAGAGAAAGCACTGTATGTTTGAAGCTGCGGAG 581

RESULT 14
BP310011
LOCUS BP310011 Sugano cDNA library, brain Homo sapiens cDNA clone
DEFINITION BP310011 Sugano cDNA library, brain Homo sapiens cDNA clone
ACCESSION BP310011
VERSION BP310011.1 GI:52238986
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 584)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1 .584
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NR01274"
/tissue type="brain"
/clone_lib="Sugano cDNA library, brain"

ORIGIN
Query Match 35.9%; Score 579.2; DB 5; Length 584;
Best Local Similarity 99.5%; Pred. No. 1.4e-139;
```

```
Matches 581; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 415 CCGTTTGAGGAAAAGGCAAGCAAGCGGATATATGAAGAAAGTGGCACCCTGTTTC 474
Db 1 CCGTTTGAGGAAAAGGCAAGCGGCGGATATATGAAGAAAGTGGCACCCTGTTTC 60
QY 475 ACCTTTAGCCTCCATCACCACCTGACATCATCGTTTACTGAAACATGCTTAATCAGGCCCAAGGAG 534
Db 61 ACCTTTAGCCTCCATCACCACCTGACATCATCGTTTACTGAAACATGCTTAATCAGGCCCAAGGAG 120
QY 535 ACTCTGTATGAGATTAAACATAGACAATAACGACGCGCATCGTCTGTGTCGGCGGAGATGGT 594
Db 121 ACTCTGTATGAGATTAAACATAGACAATAACGACGCGCATCGTCTGTGTCGGCGGAGATGGT 180
QY 595 ATGTTTCAGGAGGCTGTCACCGCTCTGATTGGGAGGACGACAGAGAGCGCCGGGTCGAC 654
Db 181 ATGTTTCAGGAGGCTGTCACCGCTCTGATTGGGAGGACGACAGAGAGCGCCGGGTCGAC 240
QY 655 CAGAACACCCCGCGGCTGTGTCGTCGCCAGTACGCTCCGATTTGGAATCATTTCCCGCA 714
Db 241 CAGAACACCCCGCGGCTGTGTCGTCGCCAGTACGCTCCGATTTGGAATCATTTCCCGCA 300
QY 715 GGGTCAACGGACTGCGTGTGTTTACTCCACCGTGGGCACCGACGCGCAGAAACCTTCGGCG 774
Db 301 GGGTCAACGGACTGCGTGTGTTTACTCCACCGTGGGCACCGACGCGCAGAAACCTTCGGCG 360
QY 775 CTGCATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTCTCAGTCCACCAACAGC 834
Db 361 CTGCATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTCTCAGTCCACCAACAGC 420
QY 835 ACACCTCTCGTACTCCGTCGCTCCGTCGTCGGCTACGGCTTCTACGGGACATCATCAAG 894
Db 421 ACACCTCTCGTACTCCGTCGCTCCGTCGTCGGCTACGGCTTCTACGGGACATCATCAAG 480
QY 895 GACAGTGAGAAAGAAACGGTGGTGGTCTTTCAGATACGACTTTTTCAGGTTTAAAGACC 954
Db 481 GACAGTGAGAAAGAAACGGTGGTGGTCTTTCAGATACGACTTTTTCAGGTTTAAAGACC 540
QY 955 TTCCTCTCCACCTGCTATGAAGGACAGTGTCTTCTCTCC 998
Db 541 TTCCTCTCCACCTGCTATGAAGGACAGTGTCTTCTCTCC 584
```

RESULT 15
BP309990
LOCUS
DEFINITION BP309990 Sugano cDNA library, brain Homo sapiens cDNA clone
NRR01216, mRNA sequence.

ACCESSION BP309990
VERSION BP309990.1 GI:52238965
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 584)
AUTHORS Suzuki.Y., Yamashita.R., Shiota.M., Sakakibara.Y., Chiba,J.,
Mizushima-Sugano.J., Nakai.K. and Sugano,S.

TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

EMAIL Email: ysuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..584

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NRR01216"
/tissue_type="brain"

ORIGIN /clone_lib="Sugano cDNA library, brain"

Query Match 35.3%; Score 570.4; DB 5; Length 584;
Best Local Similarity 99.5%; Pred. No. 2.7e-137;
Matches 582; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
QY 415 CCGTTTGAGGAAAAGGCAAGCGGATATATGAAGAAAGTGGCACCCTGTTTC 474
Db 1 CCGTTTGAGGAAAAGGCAAGCGGATATATGAAGAAAGTGGCACCCTGTTTC 59
QY 475 ACCTTTAGCCTCCATCACCACCTGACATCATCGTTTACTGAAACATGCTTAATCAGGCCCAAGGAG 534
Db 60 ACCTTTAGCCTCCATCACCACCTGACATCATCGTTTACTGAAACATGCTTAATCAGGCCCAAGGAG 119
QY 535 ACTCTGTATGAGATTAAACATAGACAATAACGACGCGCATCGTCTGTGTCGGCGGAGATGGT 594
Db 120 ACTCTGTATGAGATTAAACATAGACAATAACGACGCGCATCGTCTGTGTCGGCGGAGATGGT 179
QY 595 ATGTTTCAGGAGGCTGTCACCGCTCTGATTGGGAGGACGACAGAGAGCGCCGGGTCGAC 654
Db 180 ATGTTTCAGGAGGCTGTCACCGCTCTGATTGGGAGGACGACAGAGAGCGCCGGGTCGAC 239
QY 655 CAGAACACCCCGCGGCTGTGTCGCCAGTACGCTCCGATTTGGAATCATTTCCCGCA 714
Db 240 CAGAACACCCCGCGGCTGTGTCGCCAGTACGCTCCGATTTGGAATCATTTCCCGCA 299
QY 715 GGGTCAACGGACTCGTGTGTTTACTCCACCGTGGGCACCGACGCGCAGAAACCTTCGGCG 774
Db 300 GGGTCAACGGACTCGTGTGTTTACTCCACCGTGGGCACCGACGCGCAGAAACCTTCGGCG 359
QY 775 CTGCATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTCTCAGTCCACCAACAGC 834
Db 360 CTGCATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTCTCAGTCCACCAACAGC 419
QY 835 ACACCTCTCGTACTCCGTCGCTCCGCTTCTACGGCTTCTACGGGACATCATCAAG 894
Db 420 ACACCTCTCGTACTCCGTCGCTCCGCTTCTACGGCTTCTACGGGACATCATCAAG 479
QY 895 GACAGTGAGAAAGAAACGGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 954
Db 480 GACAGTGAGAAAGAAACGGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 539
QY 955 TTCCTCTCCACCTGCTATGAAGGACAGTGTCTTCTCTCCCT 999
Db 540 TTCCTCTCCACCTGCTATGAAGGACAGTGTCTTCTCTCCCT 584
```

Search completed: September 6, 2005, 02:19:47
Job time : 5907.4 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 5, 2005, 19:28:44 ; Search time 184.39 Seconds
(without alignments)
2892.921 Million cell updates/sec

Title: US-10-631-958-2
Perfect score: 1717
Sequence: 1 PKHLVFINPGGKCGKRI

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spo01/US10631958/runat 02092005 165811 4432/app query.fasta_1.1941
-DB=Issued Patents NA -QFMT=fastap -SUPFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CGN 1.1 397 @runat 02092005 165811 4432 -NCFU=6 -ICFU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:
3: /cgn2_6/ptodata/1/ina/5A COMB.seq:
4: /cgn2_6/ptodata/1/ina/5B COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1633.5	95.1	4432	4	US-09-774-528-148
2	613	35.7	2064	4	Sequence 148, App
3	566	33.0	1084	4	Sequence 14306, A
4	243	14.2	1857	4	Sequence 15155, A
5	243	14.2	2280	4	Sequence 3, Appli
6	237.5	13.8	2698	4	Sequence 13, Appl
7	228	13.3	1783	4	Sequence 11, Appl
8	227.5	13.2	1155	4	Sequence 1155, Ap
9	227.5	13.2	1205	4	Sequence 1, Appli
10	226	13.2	1149	4	Sequence 1, Appli
11	220.5	12.8	1533	4	Sequence 5, Appli
12	216	12.6	1611	4	Sequence 970, Appl

13	181.5	10.6	1875	4	US-09-614-221A-399
14	173.5	10.1	2462	4	Sequence 399, App
15	172	10.0	901	4	Sequence 796, App
16	171.5	10.0	1764	4	Sequence 30448, A
17	135.5	7.9	1012	4	Sequence 3940, Ap
18	134.5	7.8	1239	4	Sequence 12677, A
19	132	7.7	942	4	Sequence 263, App
20	125	7.3	485	4	Sequence 3500, Ap
21	125	7.3	885	4	Sequence 31476, A
22	125	7.3	948	4	Sequence 1079, Ap
23	118.5	6.9	3089	4	Sequence 518, App
24	117	6.8	915	4	Sequence 1861, Ap
25	116.5	6.8	900	3	Sequence 332, App
26	110	6.4	912	4	Sequence 1479, Ap
27	110	6.4	4411529	3	Sequence 1, Appli
28	109	6.3	498	4	Sequence 59, Appl
29	109	6.3	978	4	Sequence 3132, Ap
30	106.5	6.2	7195	4	Sequence 12897, A
31	106.5	6.2	7198	4	Sequence 15682, A
32	106	6.2	951	4	Sequence 1349, Ap
33	106	6.2	975	3	Sequence 1937, Ap
34	106	6.2	3761	4	Sequence 3656, Ap
35	105	6.1	13542	4	Sequence 408, App
36	105	6.1	13542	4	Sequence 154, App
37	105	6.1	13542	4	Sequence 154, App
38	101	5.9	1393	4	Sequence 40, Appli
39	100.5	5.9	4403765	3	Sequence 2, Appli
40	100	5.8	1003	3	Sequence 1064, Ap
41	98	5.7	3342	4	Sequence 748, Ap
42	96	5.6	2085	4	Sequence 1002, Ap
43	95.5	5.6	5381	3	Sequence 4, Appli
44	95.5	5.6	81001	3	Sequence 1, Appli
45	95	5.5	191	4	Sequence 28485, A

ALIGNMENTS

RESULT 1
US-09-774-528-148
; Sequence 148, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_Fl_genes Version 2.0
; SEQ ID NO 148
; LENGTH: 4432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129) .. (2817)
US-09-774-528-148
Alignment Scores:

```
Pred. No.: 8.62e-208 Length: 4432
Score: 1633.50 Matches: 313
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.01% Mismatches: 1
Query Match: 95.14% Indels: 11
DB: 4 Gaps: 1

US-10-631-958-2 (1-326) x US-09-774-528-148 (1-4432)

Qy 1 ProlysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 1591 CCAAAGCATTTACTGGTATTTATCAACCGTTTGGAGAAAGAACAGGCAACCGGATA 1650

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 1651 TATGAAGAAAGTGGACCACTGTTACCTTACCTTACCTCCATCAGCACTGATCATC--- 1707

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 1708 -----GTTACTGAACATGCTAATCAGGCCCAAGGAG 1737

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyAspGly 80
Db 1738 ACTCTGTATGAGATTAAACATAGACAAATPACACGGCATCGTCTGTGTGGCGGAGATGGT 1797

Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 1798 ATGTTACGAGAGTGTGTGACGGTCTGNTTGGAGGACGACAGAGCGCGGGGTGCAC 1857

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db 1858 CAGAACACCCCGGGCTGTGTGTCGCCAGTAGCTCCGATTCGGAATCATTCGCCGA 1917

Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 1918 GGGTCAACGGACTCGTGTGTATCTCCACCGTGGGACACGACGACGAGAACCTCGGCG 1977

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 1978 CTGCATATCGTTGTGGGACTCGTGGCCATGATGTGTCTCAGTCCACCAACACAGC 2037

Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLys 180
Db 2038 ACACCTCTTCCTACTCGTGTCTCGTGGGTACGGCTTCTACGGGACATCATCAAG 2097

Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 2098 GACAGTGAGAGAAACGGTGTGTGGTCTTCCAGATACAGACTTTTCAGGTTTAAAGACC 2157

Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 2158 TTCCTCTCCACCACTGTATGGAGGACAGTGTCTCTCTCTCCCTGCACACACACCGGTG 2217

Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 2218 GGATCTCCAAAGGATAGAGAGCCCTGCGGGCAGAGTGTCTTGTGTGGCAAGCAAG 2277

Qy 241 GlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 2278 CAGCAGCTGGAGGAGGACGACGAGAAAGCATGTATGTTTGGAGCTTGGCGGAGCGGTG 2337

Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 2338 GAGGAGTGGCAAGTCTGTGGGAGATTTCTGGCCATCAATGCCACAAACATGTCCTGT 2397

Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 2398 GCTTGTCCCGGAGGCCCGCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 2457

Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 2458 GACCTCATCTCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 2517
```

```
Qy 321 ThrAsnGlnGlnAspGln 326
Db 2518 ACCAACCCAGCAGGACCAG 2535

RESULT 2
US-09-270-767-14306
; Sequence 14306, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14306
; LENGTH: 2064
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14306

Alignment Scores: 3.06e-71 Length: 2064
Pred. No.: 613.00 Matches: 134
Score: 50.66% Conservative: 59
Percent Similarity: 35.17% Mismatches: 108
Best Local Similarity: 35.70% Indels: 80
Query Match: 4 Gaps: 7
DB:

US-10-631-958-2 (1-326) x US-09-270-767-14306 (1-2064)

Qy 2 LysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyr 21
Db 610 CGCGCCCTTTGGTCTTTATAACCCCTATGGAGTTCGCAAGCGGGAGCTCAGACCTAT 669

Qy 22 GluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGlyAsn 41
Db 670 GAGCGCCATGTGAGACCTATTTCCAGCTTTCGCGGCTAGACGCCACGTGTATC----- 723

Qy 42 LysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThr 61
Db 724 -----ACCACCTCAGAGGGCAACCAAGTGAAGGACATA 756

Qy 62 LeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyWet 81
Db 757 CTTCTGAGGCATGATCTGGAGATATACGATCGGTTTGTCTGTCGGAGCGGATGGCACC 816

Qy 82 PheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGln 101
Db 817 GTAGCAGAGGTCTCAACGGACTGATATTCGTCMAATCGAGAGTTGGGACTGGACGAA 876

Qy 102 AsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGly 121
Db 877 CAGCGGCCACCATACATTCCAAGACCG--GCTCTGCCAGTGGGTGTGATTCCCGCTGGC 933

Qy 122 SerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeu 141
Db 934 AGCACCCAGACCATTCGTTATAGTATGCACGGCAGCGGATGTGAGGACAGCGGCTATC 993

Qy 142 HisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThr 161
Db 994 CATGTGATTCGGGCCAGCATCGGGATTGGATTGTGTGCAGTGTGAGCAATGGGCACTCC 1053

Qy 162 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAsp 181
Db 1054 CTGCTCAGATCTGTGCGACAGTCTCTGAGCTACGGGTACCTGGCGGATGTGGCAGCCAG 1113

Qy 182 SerGluLysLysArgTyrPheLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 201
Db 1114 AGCGAAGAACTACCGCTGGATGGGACCGCGCGGTACAGTACAGTGGCGGTCAAGCCCTTC 1173

Qy 202 LeuSerHisHisCysTyrGluGlyThrValSerPheLeu----- 214
```



```
Db      1174 CTGAATAATCGCGCTATGACCCGCACTGAGAAATTGTAAGAGACCCGATCTTCTACTG 1233
Qy      215 -----ProAlaGlnHisThrValGlySerProArgAsp----- 225
Db      1234 ACCAGCCGCTCGAGGACATTCGCGAGAGTCCGGATAGTGTGCTCGTGGGAGAGTCA 1293
Qy      226 ---ArgLysProCysArgAlaGlyCysPheValCys----- 236
Db      1294 GTGCCATCGCTGCTATGCCCAATTGCCAGCCCTGCAGCTTCGCCAGCAGCATACAGGAA 1353
Qy      237 ArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrglyLeuGluAla 256
Db      1354 CAGCGATCCTCATTGTTTCATCAAGAGGAACTCAAGAGGCA-----GAGCGC 1401
Qy      257 AlaGluAspValGlu----- 261
Db      1402 AATCAGCAGGTAGAAACACAGAGGACTCTCATCTAGCGCCAGTGAAGCAGCACTCCTGAGG 1461
Qy      261 ----- 261
Db      1462 CCTCGTCGCGCTCCAGGCAATCTTCGATGTCCTGCCACTGGCTCCATTTCATCAATGAGGAAC 1521
Qy      262 -----GluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsn 277
Db      1522 CTCGCGCAACCATCAGTGGAGAGTTGTGGGGCAATTCTTTATGATCTGCGGCGCGAAC 1581
Qy      278 MetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp 297
Db      1582 ATAACTCGCCCTCGCGCCAGGAGTCCCAATGGCATCTCCGTTACAGTCATCTGGGTGAT 1641
Qy      298 GlySerSerAspLeuIleuIleArgLysCysSerArgPheAsnPheLeuArgPheLeu 317
Db      1642 GGTTCGCTGGACCTGATTCTCGTGAAGAACCTCACTGCTCAACACGCTGCGTTTCTG 1701
Qy      318 Ile 318
Db      1702 CTC 1704
```

RESULT 3

```
US-09-270-767-15155/c
; Sequence 15155, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15155
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15155
```

Alignment Scores:

Pred. No.:	2e-65	Length:	1084
Score:	566.00	Matches:	124
Percent Similarity:	49.58%	Conservative:	55
Best Local Similarity:	34.35%	Mismatches:	102
Query Match:	32.96%	Indels:	80
DB:	4	Gaps:	7

US-10-631-958-2 (1-326) x US-09-270-767-15155 (1-1084)

```
Qy      2 LysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTy 21
Db      1038 CGCCGCTTTTGGTCTTTTATAAACCCCTATGGAGTCCGCAAGCGGGAGCTCAGACCTAT 979
Qy      22 GluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGlyAsn 41
```

RESULT 4

```
US-09-970-516-3
; Sequence 3, Application US/09970516
```

```
Db      978 GAGCGCCATGTGAGACCTATTTTCCAGCTTCCGCGGTAGACGCCACCGTGTATC----- 925
Qy      42 LysPheTy rValAsnTy rValGluValIleThrGluHisAlaAsnGlnAlaLysGluThr 61
Db      924 -----ACCACCTCAGAGGCAAAACCAAGTGAAGGACATA 892
Qy      62 LeuTy rGluIleAsnIleAspLysTy rAspGlyIleValCysValGlyGlyAspGlyMet 81
Db      891 CTTCTGAGCCATGATCTGGGAGTATACGATCGGTTTCTGTGTGCGAGGCGATGGCACC 832
Qy      82 PheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGln 101
Db      831 GTAGCAGAGGTCAATCAACGAGTATATTCGTCAAATGCGAGAGTTGGGACTGGACGAA 772
Qy      102 AsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProIleGly 121
Db      771 CAGCGCGCCACCATCATCTCCAAAGACCG---GCTCTGCCAGTGGGTGTGATTCCTCCGCTGCG 715
Qy      122 SerThrAspCysValCysTy rSerThrValGlyThrSerAspAlaGluThrSerAlaLeu 141
Db      714 AGCACCGACACCATTTGCTATAGTAGTCACGCGCAGCGGATGTGAGGACAGCGGCTATC 655
Qy      142 HisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThr 161
Db      654 CATGTGATTCGGGCCAGCATCGGGGATGTGATGTGTCAGTGTGAGCAATGGCCAGTCC 595
Qy      162 LeuLeuArgTy rSerValSerLeuLeuGlyTy rGlyPheTy rGlyAspIleIleLysAsp 181
Db      594 CTGCTCAGATTCTGTCCAGTGTCTGAGCTACCGGTACCTCGGGCGATGTGGCAGCCAG 535
Qy      182 SerGluLysLysArgTy rPheLeuAlaArgTy rAspPheSerGlyLeuLysThrPhe 201
Db      534 AGCAGAGAACTACCCTGGATGGGACCGCGCGGTACGAGTACAGTGGCGGTCAAGGCCCTTC 475
Qy      202 LeuSerHisCysTy rGlyTy rThrValSerPheLeu----- 214
Db      474 CTGAATAATCGCGCTATGACCGCAACTGAGATGTTAGAGAGCCCGATCTTCTACTG 415
Qy      215 -----ProAlaGlnHisThrValGlySerProArgAsp----- 225
Db      414 ACCAGCGCGCTGGAGACATTCGCGAGAGTCCGGATAGTGTGCTCGCTGGGAGAGTCA 355
Qy      226 ---ArgLysProCysArgAlaGlyCysPheValCys----- 236
Db      354 GTGCCATCGCTGCTATGCCAANTTGCAGCGCTGCAGCTTCGCCAGCAGCATACAGGAA 295
Qy      237 ArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTy rGlyLeuGluAla 256
Db      294 CAGCGATCCTCATTTGTTTCATCCAGAGGAACTCAAGAGGCA-----GAGCGC 247
Qy      257 AlaGluAspValGlu----- 261
Db      246 AATCAGCAGGTAGAAACACAGAGGACTCTCATCTAGCGCCAGTGAAGCAGCACTCCTTAGG 187
Qy      261 ----- 261
Db      186 CTTGCTCGCGCTCCAGGCAATCTTCGATGTCGCCACTGGCTCCATTTTCATCAATGAGGAAC 127
Qy      262 -----GluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsn 277
Db      126 CTCGGCAACGATCAGTGGAGAGTTGTGGGGGCAATTTCTTTATGATCTGGCGGCGGAAC 67
Qy      278 MetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp 297
Db      66 ATAACCTGGCGCTGGCGCAGGAGTCCCAATGGCATCTCCCGTTACAGTCATCTGGGTGAT 7
Qy      298 Gly 298
Db      6 GGT 4
```

```
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1857)
; OTHER INFORMATION:
US-09-970-516-3

Alignment Scores:
Pred. No.: 1,34e-21 Length: 1857
Score: 243.00 Matches: 67
Percent Similarity: 47.41% Conservative: 43
Best Local Similarity: 28.88% Mismatches: 90
Query Match: 14.15% Indels: 32
DB: 4 Gaps: 4

US-10-631-958-2 (1-326) x US-09-970-516-3 (1-1857)
Qy 1 ProLysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 20
Db 430 CCGCCCCGGTGTCTTATTGGTCAATCCCTTTGGGGGTGGGGCTGGCCCTGGCAGTGG 489
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 490 TGTAAAGAACCAACGCTGCTCCCATGATCTCTGAAGCTGGGCTGTCTTCAACCTCATCCAG 549
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 550 -----ACAGAACGACAGAACCCAGCCCGGGAG 576
Qy 61 ThrLeuTyrGluIleAsnLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 577 CTGGTCCAGGGGCTGAGCCTGAGTGGGATGCGATCGTCACGGTCTCGGAGACGGG 636
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValasp 100
Db 637 CTGCTCCATGAGGTGCTGAACGGGCTCTAGATCGC----- 672
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 673 CTGACTCGGAGGAGCTGTGAAGATGCT-----GTGGGCACTCTCCCTGCG 720
Qy 121 GlySerThrAspCysValCysTyrSer-----ThrVal 131
Db 721 GGCTCGGCAACGCGCTGCGCGGACGAGTGAACAGCAGCGGGGATTTTGAGCCAGCCCTG 780
Qy 132 GlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMet 151
Db 781 GGCTCGACCTGTTGCTCAACTGCTCACTGTGCTGTCGCGGGGTGGTGGCCACCCACTG 840
Qy 152 AspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGly 171
Db 841 GACCTGTCTCTCGTGAACGCTGCGGCTCGGGCTCCCGCTGTTCTCTCTCTCTGTGTGGCC 900
Qy 172 TyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTyrTrpLeuGlyLeuAla 191
Db 901 TGGGGCTTCGTGTCAAGTGTGATATCCAGAGCGAGCGCTTCAGGGCTTGGGAGTGGCC 960
Qy 192 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrVal 211
Db 961 CGCTTCACTGGGACGCGTGTGGGCTCGCCACACTGCACCTACCGCGGAGCGCTC 1020
```

```
Qy 212 SerPheLeuProAlaGlnHisThrValGlySerPro 223
Db 1021 TCCTACCTCCCGCCCACTGTGGAAACCTGCTCGCC 1056
```

RESULT 5

```
US-09-817-676A-13
; Sequence 13, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1860)
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245447
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-13
```

```
Alignment Scores:
Pred. No.: 2,09e-21 Length: 2380
Score: 243.00 Matches: 67
Percent Similarity: 47.41% Conservative: 43
Best Local Similarity: 28.88% Mismatches: 90
Query Match: 14.15% Indels: 32
DB: 4 Gaps: 4
```

US-10-631-958-2 (1-326) x US-09-817-676A-13 (1-2380)

```
Qy 1 ProLysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 20
Db 436 CCGCCCCGGTGTCTTATTGGTCAATCCCTTTGGGGGTGGGGCTGGCCCTGGCAGTGG 495
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 496 TGTAAAGAACCAACGCTGCTCCCATGATCTCTGAAGCTGGGCTGTCTTCAACCTCATCCAG 555
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 556 -----ACAGAACGACAGAACCCAGCCCGGGAG 582
Qy 61 ThrLeuTyrGluIleAsnLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 583 CTGCTCCAGGGGCTGAGCCCTGAGTGGATGGCATCGTCAGGCTCTCGGAGACGGG 642
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValasp 100
Db 643 CTGCTCCATGAGTGTCTGAACGGGCTCTAGATCGC----- 678
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleIleIleProAla 120
Db 679 CCGTACTGGGAGGAAGCTGTGAAGATGCT-----GTGGGCACTCTCCCTGCG 726
```

Qy	121	GlySerThrAspCysValCysTySer	-----ThrVal	131
			:::	:::
Db	727	GGCTCGGCAACCGCTGGCCGGAGAGTGAACAGCACGGGGATTTGAGCCAGCCCTG		786
Qy	132	GlyThrSerAspAlaGluThrSerSerAlaLeuHisIleValValGlyAspSerLeuAlaMet		151
			:::	:::
Db	787	GGCCTCGACACTGTGTCAACTGCTCACTGTGTGTCGCCGGGTGGTGGCCACCACCTG		846
Qy	152	AspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGly		171
			:::	:::
Db	847	GACCTGCTCCGAGACGCTGGCCCTCGGGCTCCGCTGTTTCTCCTTCCTGTCGTGGCC		906
Qy	172	TyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTyrLeuGlyLeuAla		191
			:::	:::
Db	907	TGGGGCTTCGTGTCAAGTGGATATCCAGAGCCAGCGCTTCAGGGCTTTGGGCAAGTGCC		966
Qy	192	ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrVal		211
			:::	:::
Db	967	CGCTTCACACTGGGCACGGTGCTGGGCTCGCCACACTGCACACTACCGGGAGCGCTC		1026
Qy	212	SerPheLeuProAlaGlnHisThrValGlySerPro		223
			:::	:::
Db	1027	TCCTACTCTCCCGCACTGTGGAAACCTGCCTCGCCC		1062

RESIST. T. 6

```

US-09-817-676A-11
; Sequence 11, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohana, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; TITLE OF INVENTION: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817, 676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11.
; LENGTH: 2698
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (387)..(2237)
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245448
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-11

```

Alignment Scores:

Pred. No.:	1.43e-20	Length:	2598
Score:	237.50	Matches:	71
Percent Similarity:	45.97%	Conservative:	43
Best Local Similarity:	28.63%	Mismatches:	94
Query Match:	13.83%	Indels:	40
DB:	4	Gaps:	5

US-10-631-958-2 (1-326) x US-09-817-676A-11 (1-2698)

Qy 4 LeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArg 23
|||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
Db 828 CTGCTCATATTGTGCAATCCCTTTGGGGCGGGGGCTGGCTGGCAGCGCTGTATGCAC 887

Qy	24	LysValAlaProLeuPheThrLeuAlaIleSerIleThrThrAspIleIleGlyAsnLysPhe	43
Db	888	CACGTGGTCCAAATGATCTCTGAAGCTGGCGTCTCTTCAACCTCATACAG-----	938
Qy	44	TyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyr	63
Db	939	-----ACAGAACGACGAGAACCATGCCCTGAGCTGGTGCGAC	974
Qy	64	GluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSer	83
Db	975	GGGTAAAGCTAGTGAAGTGGGAAGGCAATGTCACTGTGTCTGGAGACGGCGCTTTAC	1034
Qy	84	GluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHis	103
Db	1035	GAGGTGCTGAATGGGCTCCTTGATCGG-----CAGACTGG	1070
Qy	104	ProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThr	123
Db	1071	GAGGATGCCGTGGGATGCC-----ATTGGTGTCTCCCTCTGGATCGGC	1118
Qy	124	AspCysValCys-----TyrSerThrValGlyThrSer	134
Db	1119	AATCGCTAGCTGGGGCGTGAGCCATCATGCGGGTTTGAGCAGGTGTCGGTGTGAC	1178
Qy	135	AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSer	154
Db	1179	CTGTTGCTCAACTGCTCGCTCTCTCTCTGCGGTGGTGGACCATCCTCTGGACTTGCTC	1238
Qy	155	SerValHisHisAsnSerThrLeuLeuArgTyrTyrSerValSerLeuLeuGlyTyrGlyPhe	174
Db	1239	TCTGTGACGTAGCTCGGATCCCGCTGTTTTCTTCTGTGCAGTGGCGCTGGGATTC	1298
Qy	175	TyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAsp	194
Db	1299	TTGTCAAGATGGACATTCACATGAGCGCTTCAGGGCCCTGGGCGCGCTCGATTCACA	1358
Qy	195	PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrCluGlyThrValSerPheLeu	214
Db	1359	CTGGGTGCAGTGTAGGCGTGGCTCGTTGCACTACCTACCGTGGACGCCTCTCCTACCTC	1418
Qy	215	ProAlaGlnHisThrValGlySerProArg-AspArgLysProCysArgAlaGlyCysPh	234
Db	1419	CCGCT-----ACCACGAGACCAAGCCTTGCCCATCCAGGCCAC	1457
Qy	234	eValCysArgGlnSerLysGln	241
Db	1458	AGTCTGCTCGAGCCCAAGTCAG	1479

RESULT 7

```

US-09-949-016-1155
; Sequence 1155, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1155

```


Db 451 AACCTGCTCTCTGCACACGCTTCCGGGCTCGCCTCTTCTCTGTGCTCAGCCTGGCC 510
Qy 172 TyrGlyPheTyrGlyAspIlelelyaspSerGlyLysLysArgTrpLeuGlyLeuAla 191
Db 511 TGGGGCTTCATTGCTGATGTGACCTAGAGAGTGAAGATATCGCGCTCTGGGGAGATG 570
Qy 192 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrVal 211
Db 571 CGCTTCACTCTGGGACCTTCTCTGCTGCTGGCAGCGCTCGGCACCTACCGCGCGCAGCTG 630
Qy 212 SerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 228
Db 631 GCCTACTCTCCCTGTAGGAAGA---GTGGGTTCCAAGACACCTCGCTCCCCC 678

RESULT 9

US-09-959-897-1
; Sequence 1, Application US/09959897
; Patent No. 6730480
; GENERAL INFORMATION:
; APPLICANT: PITSON, Stuart M
; APPLICANT: Brian, WATTENBERG W
; APPLICANT: Pu, XIA
; APPLICANT: Richard, D'ANDREA J
; APPLICANT: Jennifer, BABLE R
; APPLICANT: Mathew, VADAS A
; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
; FILE REFERENCE: PITSON-1
; CURRENT APPLICATION NUMBER: US/09/959,897
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00457
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: AU PQ 0339
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: AU PQ 1504
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1205
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (33)..(1184)
; OTHER INFORMATION:
US-09-959-897-1

Alignment Scores:
Pred. No.: 7, 48e-20 Length: 1205
Score: 227.50 Matches: 64
Percent Similarity: 45.15% Conservative: 43
Best Local Similarity: 27.00% Mismatches: 97
Query Match: 13.25% Indels: 33
DB: 4 Gaps: 5

US-10-631-958-2 (1-326) x US-09-959-897-1 (1-1205)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlnGlyLysArgIle 20
Db 72 CCTCGCGGCTGCTGCTGCTGCTGAACCGCGCGGCGGCAAGGGCAAGCCTTGCAGCTC 131
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleleGly 40
Db 132 TTCCGGAGTCACGTGCAGCCCTTTTGGCTGAGGCTGAATC----- 173
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 174 -----TCTTCACTGCTGATGCTCACTGAGCGCGCGGAACACCGCGCGGAG 218
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValIcylGlyAspGly 80
Db 219 CTGGTGGGTCGGAGAGCTGGGCGCTGGGAGCGCTCTGGTGGTCTATGCTGGAGACGGG 278

Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 279 CTGATGCACGAGGTGGTGAACGGCTCATGAGCGGCTGACTGGGAGACCGCCATCCAG 338
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db 339 AAG-----CCCTGTGTAGCTCCACGA 362
Qy 121 GlySerThrAspCysValCys-----TyrSerThrValGly 132
Db 363 GGCCTCTGCAACGCGCTGGCAGCTTCTTGAACCATTTATGCTGGCTATGAGCAGGTACCC 422
Qy 133 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla---Met 151
Db 423 AATGAAGACCTCTCGACCACTGACGCTATTGTCTGCCCGCGCTCTGTGACCCCATG 482
Qy 152 AspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGly 171
Db 483 AACCTGCTGCTCTGCACACGCTTCGGGGCTGCGCCTTCTCTGTGCTCAGCCTGGCC 542
Qy 172 TyrGlyPheTyrGlyAspIlelelyaspSerGlyLysLysArgTrpLeuGlyLeuAla 191
Db 543 TGGGGCTTCATTGCTGATGTGACCTAGAGAGTGAAGATATCGCGCTCTGGGGAGATG 602
Qy 192 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrVal 211
Db 603 CGCTTCACTCTGGGACCTTCTCGGCTCGGAGCGCTTGGCGACTTACCGCGCGGACTG 662
Qy 212 SerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 228
Db 663 GCTTACCTCTCTGTAGGAAGA---GTGGGTTCCAAGACACCTGCTCCCTCCCC 710

RESULT 10

US-09-970-516-5
; Sequence 5, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1149)
; OTHER INFORMATION:
US-09-970-516-5

Alignment Scores:
Pred. No.: 1, 09e-19 Length: 1149
Score: 226.00 Matches: 63
Percent Similarity: 45.15% Conservative: 44
Best Local Similarity: 26.58% Mismatches: 94
Query Match: 13.16% Indels: 36
DB: 4 Gaps: 5

US-10-631-958-2 (1-326) x US-09-970-516-5 (1-1149)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlnGlyLysArgIle 20
Db 40 CCATGACAGGTGGTGGTCTGCTGAACCCCGCGGTCGCAAGGGCAAGCCTTGCAGCTC 99
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleleGly 40
Db 100 TTCCAGACCGCTGTGACGCCCTTCTTGGAGGAGGAGAGATACCTTTAAACTGATA--- 156

```
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 157 -----CTCACCGAAGCGAAGAACCATGCCAGGAG 186
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 187 CTGGTGTGTGCAGAGGAGTTGGGTCACTGGGACGCCCTGGCAGTCATGTCGGGTGATGGT 246
Qy 81 MetPheSerGluValLeuHisGlyIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 247 CTGATGATGAGGTGGTGAATGGGCTAATGAAACGCCGCTGGGAGACTGCCATCCAG 306
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 307 AAA-----CCCTGTGTAGCTCCCTGGA 330
Qy 121 GlySerThrAspCysValCys-----TyrSerThrValGly 132
Db 331 GGCTCCGCAATGCCTGGCAGCTTCTGTGAACCACTATGCTGGGTACGAGCAGGTGACT 390
Qy 133 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla--Met 151
Db 391 AATGAAGACCTGCTCATCACTGCACACTGCTGTGTGCCCGCGCCTGTACCCATG 450
Qy 152 AspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGly 171
Db 451 AACCTGTGCTCCCTGCACACTGCTTCTGGGCTGCGGCTATTCTGTGCTCAGTCTGTCC 510
Qy 172 TyrGlyPheTyrGlyAspIleIleLysAspSerGluLysArgTyrTrpLeuGlyLeuAla 191
Db 511 TGGGGCTTTGTTGTGCTGAGTCTGCAGAGTGTGAGAGTGTGAGAGTGTGGGGGAGATT 570
Qy 192 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrVal 211
Db 571 GGTTCACAGTGGGACCTCTTTCGCTCAACAGCCTGCCCATCTACCAAGGCCAACTG 630
Qy 212 SerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 228
Db 631 GCCTACCTTCT-----GTAGGAACCTGGGCTCTAAGAGACCC 669

RESULT 11
US-09-205-258-90
; Sequence 90, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007F1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
```

```
; LOCATION: (1522)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1527)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-90

Alignment Scores:
Pred. No.:      1e-18      Length:      1533
Score:          220.50     Matches:      70
Percent Similarity: 45.80%   Conservative: 39
Best Local Similarity: 29.41% Mismatches:    95
Query Match:      12.84%     Indels:      34
DB:               4         Gaps:         7

US-10-631-958-2 (1-326) x US-09-205-258-90 (1-1533)

Qy  1  ProLysHisLeuValPheIleAenProPheGlyGlyGlyGlnGlyLysArgIleThrGluArg 20
Db  112 CCTCGCGGTGCTGGTGTGCTGAACCCGCGCGCGCAAGGCAAGGCTTGAGCTC 171
Qy  21  TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db  172 TTCGGAGTCAGTCGACGCCCTTTGGCTGAGGCTGAATC----- 213
Qy  41  AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db  214 -----TCCTTCACGCTGATGCTCACTGAGCGCGCGAACCACGCGGGGAR 258
Qy  61  ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyAspGly 80
Db  259 CTGGTGGCGTCGAGGAGCTGGCGCGCTGGRACGCTCTGGTGTGTCATGTGTGGAGACGG 318
Qy  81  MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db  319 CTGATGCAGAGGTGTGACGGGCTTCATGGA-----GCCGCTGACGGGA 366
Qy  101 GlnAsnHisProArgAlaValLeuValPro-SerSerLeuArgIleGlyIleIleProAl 120
Db  367 GACCGCCATCCAGAGCC-----CCTGTGTAGCTC-----CCAGC 402
Qy  120 aGlySerThrAspCysValCys-----TyrSerThrValG 132
Db  403 AGGCTCTGCAACGCSCTGGCAGCTTCCTTRAACCATATTGCTGGCTATRAGCAGGTAC 462
Qy  132 yThrSerAspAlaGluThrSerAlaLeuHisIleValGlyAspSerLeuAla--Me 151
Db  463 CAATGAAGACCTCTGACCACTGACGCTATTGCTGCGCGCGCTGCTGTACCCAT 522
Qy  151 tAspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuG 171
Db  523 GAACCTGTCTCTGCAACAGCTTCGGGGCTGCGCTCTTCTCTGTGCTCAGCCTGCG 582
Qy  171 yTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAl 191
Db  583 CTGGGGCTTCATTGTGTGATGTGGACCTAGAGAGTGAGAAGTATCGCGCTCTGGGGGAGAT 642
Qy  191 aArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrVa 211
Db  643 GCGCTTCACTCTGCGGACCTCTCTCGCTGCGACCTGCGACCTACCGCGCGCGACT 702
Qy  211 lSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 228
Db  703 GGCCTACTCTCTGTAGGAGA--GTGGTTCCAGACACCTGCTCCCC 751

RESULT 12
US-09-248-796A-1756
; Sequence 1756, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
```

```
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1756
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Candida albicans
US-03-248-796A-1756

Alignment Scores:
Pred. No.:      4.41e-18      Length:      1611
Score:          216.00     Matches:      79
Percent Similarity: 42.95%   Conservative: 58
Best Local Similarity: 24.76% Mismatches:    131
Query Match:      12.58%     Indels:      52
DB:               4         Gaps:         10

US-10-631-958-2 (1-326) x US-09-248-796A-1756 (1-1611)

Qy  4  LeuLeuValPheIleAenProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArg 23
Db  493 ATTTTGGTATTGATAAACCCGCGCATGGCGGCCAAGCACGCAAACTATTTCACAAAAT 552
Qy  24  LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGlyAsnLysPhe 43
Db  553 AAAATCTTACCAATATTACAAAGCGCT-----CGTGCT 585
Qy  44  TyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyr 63
Db  586 AATGTTACGATTATTGAA-----ACTAAATATCATGGACACGCCACTGAGATTGCGCGT 639
Qy  64  GluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSer 83
Db  640 GAGCTAGATGCAATGATTATGATAAATTGTTGTTGTTCTGCGCATGGGATACCTCAT 699
Qy  84  GluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHis 103
Db  700 GAAGTTATCAATGGCTCTCTATCTTCGCCAGAT----- 732
Qy  104 ProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThr 123
Db  733 ---AAAGGTTTACTGGCATTCAACAAAATTCAGTTACTCAATTACCTTGTGGGTGAGG 789
Qy  124 AspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIle 143
Db  790 AACGCGTTGAGTTTGAGTACACATGTTAGTAAAAATGCTTCAGTTGCAACTCTTTATATG 849
Qy  144 ValValGlyAspSerLeuAlaMetAspValSerSerValHisHis-----AsnSer 160
Db  850 TTGAAGCTCATAGACAAAATTCGATTGTAGTCTATTACCCAAAGGTACAGGAAGTAA 909
Qy  161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLys 180
Db  910 AAGATAAGAAATTTGTCATTTTAAAGTCAGTGTACGGTATTATTGCTGATTCTGATATT 969
Qy  181 AspSerGluLysLysArgTrpLeuAlaArgTyrAsp----- 194
Db  970 GGAACAGACAACTTTACGTTGGTGGTCCCTATACGGTTTGAACCTGGAGTGATACAAAA 1029
Qy  195 ---PheSerGlyLeuLys-----ThrPheLeuSerHisHisCysTyrGluGly 209
Db  1030 GTGTCTCTCGGGGCAAAATATCTCTGTGACCTATTGTTGAAATACAAG---TACGATAAT 1086
Qy  210 ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCys 229
Db  1087 AATTCAGAGATTTTGAATCATGTAAATGATTACTTGTAGTAATATGAT----- 1134
```

Qy 230 ArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLysLys 249
Db 1135 -----ACTGAAACAGTGTACCATTTGTCTGACGAGCAACTTG 1173
Qy 250 AlaLeuTyrGlyLeuGluAlaAlaGluaspVal---GluGluTTPGln-ValValCysG1 268
Db 1174 CAATAACGAGCCCTGATTATAGATCAACCTGTCCCAATGATTGGAGACATATTCCTCAA 1233
Qy 268 YLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArg-- 287
Db 1234 GAAATTTCTATAATTTGAACATCTTGTATGTTGGCAAAATGCCATTCGTGTCGAGAT 1293
Qy 288 ----GlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeu 304
Db 1294 ACTCAGTTTTTC-CCTGCTGCTCTACCGAAGCATGGTTCAATGGACATGATTGTC 1347

RESULT 13

US-09-614-221A-399
; Sequence 399, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; FILE REFERENCE: WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 399
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-399

Alignment Scores:
Pred. No.: 2,51e-13 Length: 1875
Score: 181.50 Matches: 62
Percent Similarity: 42.68% Conservative: 43
Best Local Similarity: 25.20% Mismatches: 96
Query Match: 10.57% Indels: 45
DB: 4 Gaps: 7

US-10-631-958-2 (1-326) x US-09-614-221A-399 (1-1875)

Qy 2 LysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyr 21
Db 679 AGATCGATATTAGTCATTATTAAATCCCAACGGTGTAAAGTACTGCTAAAATTTATTC 738
Qy 22 GluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGlyAsn 41
Db 739 CTGACAAAAGCAAGCCCAATACTAGTGGAAAGT-----GGCTGC 777
Qy 42 LysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThr 61
Db 778 AAAATAGAAATTCATAC-----ACAAAATATGCCGTCACGCCATCGATATT 825
Qy 62 LeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMet 81
Db 826 GCCAAAGATTATAGATATCAGCAATACGATACCATTCATGCTGCTGGGTGATGTTATT 885
Qy 82 PheSerGluValLeuHisGlyIleGlyArgThrGlnArgSerAlaGlyValAspGln 101
Db 886 CCATACGAAGTAATTAATGGCTTTATAGAAGACCCGACAGA-----GTGGATGCG 936
Qy 102 AsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGly 121
Db 937 TTCAATAAACATAGCGGTAACTACG-----CTACCTTCGGGT 972
Qy 122 SerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeu 141

Db 973 TCAGGAAATGCTATGAGCATTTTCATGTTCATTGGACAAATAAACCCATCGTACGCGCTCTG 1032
Qy 142 HisIleValValGlyAspSerLeuAlaMetAspVal-----SerSerValHisHis 158
Db 1033 TGCCTTGTCAAAATCCATTGAAACAAGAAATAGACTTAATGTGTGTTCCCAACCTTCTTAT 1092
Qy 159 AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 178
Db 1093 ATGAACGAATGGCCCAAGATTATCTTTTTCAGTCAGACGGTACGGCGTTATTGCAGAATCT 1152
Qy 179 IleLysAspSerGluLysLysArgTTPLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 198
Db 1153 GATATTAAACACTGAATTCATCAGATGGATGGTCCCGTCCCTAGGTTTAATTTGGGT----- 1206
Qy 199 LysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 218
Db 1207 -----GTAGCATTCACATTTATCCAAAGT----- 1230
Qy 219 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 238
Db 1231 -----AGAAATATCCCTGCGAAGTTTTCGTCAAATATGCTGCCCAA 1272
Qy 239 SerLysGlnGlnLeuGlu 244
Db 1273 TCTAAAAAGGAATTAATA 1290

RESULT 14

US-09-620-312D-796
; Sequence 796, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polyptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 796
; LENGTH: 2462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (196)..(1464)
US-09-620-312D-796
Alignment Scores:
Pred. No.: 4.85e-12 Length: 2462
Score: 173.50 Matches: 62
Percent Similarity: 45.38% Conservative: 46


```
Best Local Similarity: 26.05% Mismatches: 83
Query Match: 10.10% Indels: 47
DB: 4 Gaps: 11

US-10-631-958-2 (1-326) x US-09-620-312D-796 (1-2462)

QY 2 LysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyr 21
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
376 AAGAGGCCACTGTTTCTCAATCTGCAGCTTGCNAGGAAAGCCAGGACTATTT 435
QY 22 GluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGlyAsn 41
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
436 GAAAAAATGCTGCCCGATTTTACATTTATCTGCGCATGGATGTGACTATT 486
QY 42 LysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThr 61
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
487 -----GTTAAGACAGATTATGAGGGCAAGCCCAAGAAA--- 519
QY 62 LeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMet 81
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 CTCCTGGAACG---ATGGAACACGGATGTGATCATTTGTCAGGAGGATGGACA 576
QY 82 PheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGln 101
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
577 CTGAGGAGTTGTTACTGTCTTCGACGACAGATGAGGCTACCTTC----- 627
QY 102 AsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGly 121
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
628 -----AGTAAGATTCCCATTTGGATTTATCCCACTGGGA 660
QY 122 SerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr----- 138
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 GAGACAGTAGTTGAGTCATACCTCTTTGCGAAGTGGAAACAAAGTCCAAATATT 720
QY 139 -----SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerVal 156
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
721 ACTGATGCCACACTTGCCATTGTGAAGAGAGACAGTCCACTTGATGTCTTGAGATC 780
QY 157 HisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGly 176
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
781 AAGGGTGAAGAGAACAGACCTGTATTGCAATGACCGGCTTCGATGGGATCTTTTCA 840
QY 177 AspIle---IleLysAspSerGluLysLysArgTyrP-----LeuGlyLeu--- 190
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
841 GATGTCGGCTCAAGATTAGC-----AAGTACTGGTATCTTGAGCCTCTTAAATAACAA 894
QY 191 AlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGlyThr 210
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
895 GCAGCCCACTTTTTCAGCACTCTTAAGGAGTGGCCTCAGACTCAT-----CAAGCCTCT 948
QY 211 ValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 228
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
949 ATCTCATACACG-----GGACCTTACAGAGAGACCTCCC 981

RESULT 15
US-09-270-767-30448
; Sequence 30448, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30448
; LENGTH: 901
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-30448
Alignment Scores:

Pred. No.: 1.28e-12 Length: 901
Score: 172.00 Matches: 50
Percent Similarity: 38.59% Conservative: 21
Best Local Similarity: 27.17% Mismatches: 45
Query Match: 10.02% Indels: 68
DB: 4 Gaps: 5

US-10-631-958-2 (1-326) x US-09-270-767-30448 (1-901)

QY 199 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu----- 214
DB 2 AAGGCTTCTCTGAATAATCGCGCTATGACGCCGAATCTGAGAATGTTAGAGAGCCGAT 61
QY 215 -----ProAlaGlnHisThrValGlySerProArgAsp----- 225
DB 62 CTTTACTGACACCGCGCTGGAGACATTCGCGAGATCCGGATAGTGTGTGCTCGTGTG 121
QY 226 -----ArgLysProCysArgAlaGlyCysPheValCys----- 236
DB 122 GGAGAGTCAGTGCCATCCGCTCTGTATGCCAATTCGCGAGCGCTGCGAGCTTCGCCAGCAGC 181
QY 237 -----ArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrGly 253
DB 182 ATACAGGAACAGCGATCTCATTTGTTTCATCAAGAGGAATCTAAAGAGGCA----- 232
QY 254 LeuGluAlaAlaGluAspValGlu----- 261
DB 233 ---GAGCGCAATCAGCAGGTAGAACAGAGGACTCTCATCTAGCCGCCAGTGAAGACGA 289
QY 261 ----- 261
DB 290 CTCCTGAGGCTCTGTCGCGCTCCAGGCAATCTTCGATTGCCACTGGCTCCATTTCATCA 349
QY 262 -----GluTrpGlnValValCysGlyLysPheLeuAlaIleAsn 274
DB 350 ATGAGGAACCTCGGCAACGATCAGTGGAAAGTTGTGCGGGGCAATTTCTTTATGATCTGC 409
QY 275 AlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaIleHis 294
DB 410 GCGCGAACAATAACCTGCGCTGCGCCAGGAGTCCCAATGGCATCTCCCGTTTACAGTCAT 469
QY 295 LeuGlyAspGlySerSerAspLeuIleLeuArgLysCysSerArgPheAsnPheLeu 314
DB 470 CTGGTGATGTTGCTGGACCTGATTCTCGTGAAGAAAGACCTCACTGCTCTCAACACGTG 529
QY 315 ArgPheLeuIle 318
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
530 CGTTTTCTGCTC 541

Search completed: September 6, 2005, 02:33:28
Job time : 208.39 secs
```

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: September 5, 2005, 23:41:54 ; Search time 918.291 Seconds
(without alignments)
2324.913 Million cell updates/sec

Title: US-10-631-958-2
Perfect score: 1717
Sequence: 1 PKHLLVFNPFGKQKQKRI.....KCSRNFELRLRHTNQDQ 326

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7338684 seqs, 3274456166 residues
Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-DB=US10-631-958-2 -O=FASTA -SUFFIX=us10-631-958-2 -MINMATCH=0.1
-LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10631958 @CGN 1 1588 @runat 02092005 165813 4470
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPEXT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1717	100.0	979	10	US-09-969-896-1	Sequence 1, Appli
2	1717	100.0	979	19	US-10-631-958-1	Sequence 1, Appli
3	1717	100.0	1840	9	US-09-784-810A-5	Sequence 5, Appli
4	1717	100.0	1840	22	US-10-876-281-5	Sequence 5, Appli
5	1640.5	95.5	1614	10	US-09-969-896-9	Sequence 9, Appli
6	1640.5	95.5	1614	19	US-10-631-958-9	Sequence 9, Appli
7	1640.5	95.5	1740	18	US-10-632-511-39	Sequence 39, Appli
8	1640.5	95.5	4413	10	US-09-969-896-16	Sequence 16, Appli
9	1640.5	95.5	4413	19	US-10-631-958-16	Sequence 16, Appli
10	1640.5	95.5	4429	19	US-10-618-941-55	Sequence 55, Appli
11	1633.5	95.1	4432	17	US-10-120-988-148	Sequence 148, App
12	1632.5	95.1	4463	16	US-10-315-597A-1	Sequence 1, Appli
13	811	47.2	474	10	US-09-969-896-4	Sequence 4, Appli
14	811	47.2	474	19	US-10-631-958-4	Sequence 4, Appli
15	529	30.8	2657	20	US-10-425-115-176413	Sequence 176413,
16	508.5	29.6	329	10	US-09-969-896-5	Sequence 5, Appli
17	508.5	29.6	329	19	US-10-631-958-5	Sequence 5, Appli
18	432.5	25.2	2079	18	US-10-425-114-34530	Sequence 34530, A
19	432.5	25.2	2079	20	US-10-425-115-86211	Sequence 86211, A
20	414.5	24.1	564	22	US-10-477-445-46	Sequence 46, Appli
21	335	19.5	1833	19	US-10-437-963-30650	Sequence 30650, A
22	334.5	19.5	2166	20	US-10-425-115-64890	Sequence 64890, A
23	333	19.5	1774	18	US-10-425-114-34081	Sequence 34081, A
24	333	19.4	2084	18	US-10-424-599-47396	Sequence 47396, A
25	330	19.2	1869	18	US-10-425-114-34806	Sequence 34806, A
26	321.5	18.7	2189	19	US-10-437-963-92801	Sequence 92801, A
27	318	18.5	3362	20	US-10-425-115-29002	Sequence 29002, A
28	315.5	18.1	1512	20	US-10-425-115-64892	Sequence 64892, A
29	308.5	18.0	2014	18	US-10-424-599-59419	Sequence 59419, A
30	305.5	17.8	1857	18	US-10-425-114-32139	Sequence 32139, A
31	305.5	17.8	2247	20	US-10-425-115-97849	Sequence 97849, A
32	302	17.6	167	10	US-09-969-896-6	Sequence 6, Appli
33	302	17.6	167	19	US-10-631-958-6	Sequence 6, Appli
34	300.5	17.5	817	19	US-10-115-635-247	Sequence 247, App
35	300	17.5	1386	19	US-10-437-963-95148	Sequence 95148, A
36	291	16.9	2629	19	US-10-437-963-32906	Sequence 32906, A
37	289.5	16.9	1468	19	US-10-767-701-13363	Sequence 13363, A
38	288.5	16.8	2284	20	US-10-425-115-79424	Sequence 79424, A
39	287.5	16.7	3008	19	US-10-437-963-77896	Sequence 77896, A
40	282	16.4	2629	17	US-10-348-052-24	Sequence 24, Appli
41	282	16.4	2629	19	US-10-622-011-24	Sequence 24, Appli
42	276	16.1	861	20	US-10-425-115-86219	Sequence 86219, A
43	273	15.9	522	9	US-09-784-810A-7	Sequence 7, Appli
44	273	15.9	522	22	US-10-876-281-7	Sequence 25, Appli
45	270.5	15.8	2609	17	US-10-348-052-25	

ALIGNMENTS

RESULT 1
US-09-969-896-1
; Sequence 1, Application US/09969896
; Publication No. US2003012553A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-1

Alignment Scores:
Pred. No.: 1.24e-211 Length: 979
Score: 1717.00 Matches: 326
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-631-958-2 (1-326) x US-09-969-896-1 (1-979)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 2 CCAAGCATTTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACCAAGCGGATA 61
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 62 TATGAAAGAAAGTGGCACCACCTGTTACCTTAGCCTCCATCACCCTGACATCATCGGT 121
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 122 AACAAATCTATGTTAACTATGTAAGTAATTAATCAACATGCTTACCATCTGACATCATCGGT 181
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 182 ACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTCTGTGTCGGCGGAGATGGT 241
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 242 ATGTTTCAGCAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTGCAC 301
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db 302 CAGAACCAACCCCGGGCTGCTGCTGCCAGTAGCTCCGATGGAAATCATTTCCGCA 361
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 362 GGGTCAACCGACTCGTGTGTACTCCACCGTGGGACCCAGCGACGCGAGAAACCTCGCG 421
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSer 160
Db 422 CTGCATATCGTGTGTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACAGC 481
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 482 ACATCTCTCTACTCCGTCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 541
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 542 GACAGTGAGAGAAACCGGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 601
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 602 TTCCTCTCCCAACCATGCTATGAAAGGACAGTGTCTTCCTCCCTGCACACACACCGTG 661
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 662 GGATCTCCAGGATAGGAAGCCCTGCGCGCAGAGATGCTTTGTTGAGCAAGCAAGCAAG 721
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 722 CAGCAGCTGGAGGAGGACAGAGAAACACTGTATGGTTTGGAAAGCTGCGGAGACGTG 781
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 782 GAGGAGTGGCAAGTCGTGTGGGAAGTTTCTGGCCATCAATGCCACCAACATGTCCTGT 841

; SEQUENCE INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-1
; Sequence 1, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-1

Alignment Scores:
Pred. No.: 1.24e-211 Length: 979
Score: 1717.00 Matches: 326
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-631-958-2 (1-326) x US-10-631-958-1 (1-979)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 2 CCAAGCATTTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACCAAGCGGATA 61
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 62 TATGAAAGAAAGTGGCACCACCTGTTACCTTAGCCTCCATCACCCTGACATCATCGGT 121
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 122 AACAAATCTATGTTAACTATGTAAGTAATTAATCAACATGCTTACCATCTGACATCATCGGT 181
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 182 ACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTCTGTGTCGGCGGAGATGGT 241
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 242 ATGTTTCAGCAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTGCAC 301
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db 302 CAGAACCAACCCCGGGCTGCTGCTGCCAGTAGCTCCGATGGAAATCATTTCCGCA 361
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 362 GGGTCAACCGACTCGTGTGTACTCCACCGTGGGACCCAGCGACGCGAGAAACCTCGCG 421
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSer 160
Db 422 CTGCATATCGTGTGTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACAGC 481
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 482 ACATCTCTCTACTCCGTCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 541
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 542 GACAGTGAGAGAAACCGGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 601
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 602 TTCCTCTCCCAACCATGCTATGAAAGGACAGTGTCTTCCTCCCTGCACACACACCGTG 661
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 662 GGATCTCCAGGATAGGAAGCCCTGCGCGCAGAGATGCTTTGTTGAGCAAGCAAGCAAG 721
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 722 CAGCAGCTGGAGGAGGACAGAGAAACACTGTATGGTTTGGAAAGCTGCGGAGACGTG 781
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 782 GAGGAGTGGCAAGTCGTGTGGGAAGTTTCTGGCCATCAATGCCACCAACATGTCCTGT 841
```

Db 362 GGGTCAACGACTGCGTGTGTTACTCCACCGTGGGCACGACGACGAAACCTCGGCG 421
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 422 CTGCATATCGTGTGGGACTCGCTGGCGATGGATGTCTCAGTCCACCAACAGC 481
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 482 ACACATCCCTTCGCTACTCCGCTGTCCTGTGGGCTACGGCTTCTACGGGGACATCATCAAG 541
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 542 GACAGTGAGAAACGGTGTGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGACC 601
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 602 TTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTGCGCAACACACGCGTG 661
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 662 GGATCTCCAAAGGATAGGAAGCCCTGCGGGCAGATGCTTTGTTTGGAGGCAAGCAAG 721
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 722 CAGCAGCTGGAGAGAGACAGAAAGCACTGTATGGTTTGGAGCTGGGAGACGCTG 781
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 782 GAGGAGTGGCAAGTCTGTCTGGGAAGTTTCTGGCCATCAATGTCACAAACATGTCTGT 841
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 842 GCTTGTCCGCGAGAGCCCGAGGGCCCTCTCCCGGCTGCCCCACTTGGGAGACGGGTCTTCT 901
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 902 GACCTCATCTCTCATCCGGAATGCTCCAGGTTCATATTTTCTGAGATTTCTCATCATGAGCAC 961

RESULT 3

US-09-784-810A-5
; Sequence 5, Application US/09784810A
; Patent No. US20020082203A1

GENERAL INFORMATION:

; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING

; TITLE OF INVENTION: SAME

; FILE REFERENCE: 10716-08

; CURRENT APPLICATION NUMBER: US/09/784,810A

; CURRENT FILING DATE: 2001-02-14

; PRIOR APPLICATION NUMBER: 60/182,360

; PRIOR FILING DATE: 2000-02-14

; PRIOR APPLICATION NUMBER: 60/191,261

; PRIOR FILING DATE: 2000-03-22

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 1840

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-784-810A-5

Alignment Scores:

Pred. No.:	3,31e-211	Length:	1840
Score:	1717.00	Matches:	326
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-631-958-2 (1-326) x US-09-784-810A-5 (1-1840)
Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 265 CCAAAGCATTTACTGGTATTTTATCAACCCGTTTGGAGGAAAGGACAAAGGCAAGCGGATA 324
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 325 TATGAAAGAAAGTGGACCACTGTTACCTTAGCCTCCATCACCACATGACATCATCGGT 384
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 385 AACAAATTTCTATGTTAACTATGTAGAGTANTTACTGAACATGCTTAATCAGGCCAAGGAG 444
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 445 ACTCTGTATGAGATTTAAACATAGACAAATACACGCGCATCGTCTGTGTCGGGGAGATGGT 504
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 505 ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTCCGAC 564
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 565 CAGAACCAACCCCGGGCTGTCTGGTCCCAGTAGCTCCGGATTGGAAATCATTTCCCGCA 624
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 625 GGGTCAACGGACTGCGTGTGTTTACTCCACCGCTGGGCACGACGACGAGAACCTCGGCG 684
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 685 CTGCATATCGTGTGTTGGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACACAGC 744
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 745 ACATCTCTTCCTACTCGTGTCTCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 804
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 805 GACAGTGAGAAAGAACCGGTGTTGGTCTTCCAGATACGACTTTTCAAGTTTAAAGACC 864
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 865 TTCTCTCTCCACCACTGCTATGAAGGGACAGTGTCTTCTCCCTGCGCAACACACGCGTG 924
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 925 GGATCTCCAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGGAGCTGCGGAGACGAG 984
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 985 CAGCAGCTGGAGAGGAGGACAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAGACGCTG 1044
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 1045 GAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTGT 1104
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1105 GCTTGTCCCGGAGGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 1164
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1165 GACCTCATCTCATCCGGAATGCTCCAGGTTCATATTTTCTGAGATTTCTCATCATGAGCAC 1224
Qy 321 ThrAsnGlnGlnAspGln 326
Db 1225 ACCAACCAAGCAGGACCAG 1242

RESULT 4

US-10-876-281-5

; Sequence 5, Application US/10876281

```
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; PRIOR FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-876-281-5

Alignment Scores:
Pred. No.: 3,31e-211 Length: 1840
Score: 1717.00 Matches: 326
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-631-958-2 (1-326) x US-10-876-281-5 (1-1840)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 20
Db 265 CCAAAGCATTTACTGGTATTATACCCGTTTCGAGAAAGACCAAGCAAGCGGATA 324

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleGly 40
Db 325 TATGAAGAAAGTGGCACCACCTGTTACCTTAGCCTCCATCACCACCTGATCATCGGT 384

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 385 AACAAATTTCTATGTAATATGTAGAAAGTAATTAATCTGAACATGCTAATCAGGCCAAGGAG 444

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyLysValGlyGlyAspGly 80
Db 445 ACTCTGATGAGATTAACATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGGT 504

Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 505 ATGTTTCAGCGAGGTGTCACGGTCTGATTGGGAGGACGACAGAGCGCGCGGTCGAC 564

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 565 CAGAACCAACCCCGCGGTGTCGTGTCGCCAGTAGCCTCCGGATTGGAATCATTTCCCGCA 624

Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 625 GGGTCAACGGACTGGGTGTGTACTTCCACCGTGGGCACACGACGAGAAACCTCGCG 684

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSer 160
Db 685 CTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCCTCAGTCCACCAACAGC 744

Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 745 ACATCTCTTCGTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 804

Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 805 GACAGTGAGAAAGACGGTGGTGTGCTTCCAGATACGACTTTTCAGGTTTAAAGACC 864

Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220

; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; PRIOR FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-876-281-5

Alignment Scores:
Pred. No.: 3,31e-211 Length: 1840
Score: 1717.00 Matches: 326
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-631-958-2 (1-326) x US-09-969-896-9 (1-1614)

; Sequence 9, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-9

Alignment Scores:
Pred. No.: 2,22e-201 Length: 1614
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 95.54% Indels: 11
DB: 10 Gaps: 1

US-10-631-958-2 (1-326) x US-09-969-896-9 (1-1614)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 20
Db 388 CCAAAGCATTTACTGGTATTATTAACCCGTTTGGAGAAAAGGACAAAGCAAGCGGATA 447

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 448 TATGAAGAAAGTGGCACCACCTGTTCCATTCCTCCATCACCCTGATCATCATC--- 504

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 505 -----GTTACTGAACATGCTATCATCGGCCAAGGAG 534

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
```

```
Db 535 ACTCTGTATGAGATTAAACATAGACAAATACGACGGCATGCTGTGTGCGCGAGATGCT 594
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 595 ATGTTTCAGCGAGTGCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCGCGGGTCCGAC 654
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 655 CAGAACACCCCGGGCTGTGTGTCCTCCAGTAGCTCCGGATTGGAAATCATTCCTCCGCA 714
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 715 GGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGCACCCAGCGACGACGAGAACTTCGGG 774
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 775 CTGCATATCTGTTGTTGGGACTCGCTGGCCATGGATGTGCTCCTCAGTCCACCACACAGC 834
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 835 ACACCTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 894
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 895 GACAGTGAGAGAAACGGTGTGTTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACC 954
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 955 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCTCCACACACAGCGTG 1014
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 1015 GGATCTCCAGGGATAGGAAGCCCTGCGCGGACAGTGTCTTCTCTCCCTGCAACACACAGG 1074
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 1075 CAGCAGCTGGAGAGAGAGAGAAAGCAGTGTATGGTTTGGAAAGCTGCGGAGGACGCTG 1134
Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 1135 GAGGAGTGGCAAGTCTGCTGTGGGAAGTTCTGCGGCATCAATGCTTCTGAGATTTCTCATCAGG 1194
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1195 GCTTGTGCGCGAGCGCCAGGCGCTCTCCCGGCTGCGCCACTTGGGAGACGGGTCTTCT 1254
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1255 GACCTCATCTCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGAC 1314
Qy 321 ThrAsnGlnGlnAspGln 326
Db 1315 ACCAACACGACGAGACACAG 1332
```

RESULT 6

```
US-10-631-958-9
; Sequence 9, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
```

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-9

Alignment Scores:

Pred. No.: 2,22e-201 Length: 1614
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 95.54% Indels: 11
DB: 19 Gaps: 1

US-10-631-958-2 (1-326) x US-10-631-958-9 (1-1614)

```
Qy 1 ProlYsHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 388 CCAAGCAATTTACTGTTATTTATCACCCTTTGGAGGAAAGGACAGGACGCGGATA 447
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 448 TATGAAGAAAGTGGCACCCTGTTCACTTAGCTTCATCACCACCATCATCATC--- 504
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 505 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 534
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 535 ACTCTGTATGAGATTAACTAGACAAATACGACGGCATGCTGTGTGCGCGAGATGCT 594
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 595 ATGTTTCAGCGAGTGTGTCACGGTCTGATTGGAGGACGACAGAGAGCGCGGGTCCGAC 654
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 655 CAGAACACCCCGGGCTGTGTGTCCTCCAGTAGCTCCGATTCGGAATCATTCCTCCGCA 714
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 715 GGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGCACGACGACGAGAACTTCGGG 774
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 775 CTGCATATCTGTTGTTGGGACTCGCTGGCCATGGATGTGCTCCTCAGTCCACCACACAGC 834
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 835 ACACCTCTTCGCTACTCCGCTGCTCCCTGCGGCTACGGCTTCTACGGGGACATCATCAAG 894
Qy 181 AspSerGluLysLysArgTyrPheLeuGluAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 895 GACAGTGAGAGAAACGGTGTGTTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACC 954
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 955 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCAACACACAGG 1014
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 1015 GGATCTCCAGGGATAGGAAGCCCTGCGCGGACAGTGTCTTCTCTCCCTGCAACACACAGG 1074
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 1075 CAGCAGCTGGAGAGGAGGACAGAGAAAGCAGTGTATGGTTTGGAAAGCTGCGGAGGACG 1134
Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 1135 GAGGAGTGGCAAGTCTGCTGTGGGAAGTTCTGCGGCATCAATGCTTCTGAGATTTCTCATCAGG 1194
```

```
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1195 GCTTGTGCCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTCT 1254
Qy 301 AspLeuLeuLeuLeuArgLysCysSerArgPheAlaSerPheLeuArgPheLeuLeuLeuArgHis 320
Db 1255 GACCTCATCCCTCATCCCGGAATGTCACAGGTTCATTTTCTGAGATTTCTCATCAGGCAC 1314
Qy 321 ThrAsnGlnGlnAspGln 326
Db 1315 ACCAACCCAGCAGGACCAG 1332
```

RESULT 7

```
US-10-262-511-39
; Sequence 39, Application US/10262511
; Publication No. US20040038223A1
```

GENERAL INFORMATION:

```
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Chatterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
```

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

```
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
```

```
; SEQ ID NO 39
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (76)..(1686)
US-10-262-511-39
```

Alignment Scores:

```
Pred. No.: 2.5e-201 Length: 1740
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 95.54% Indels: 11
DB: 18 Gaps: 1
```

```
US-10-631-958-2 (1-326) x US-10-262-511-39 (1-1740)
```

```
Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 463 CCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAAGGACAAGCAAGCGGATA 522
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 523 TATGAAAGAAAAGTGGCACCACCTGTTCACTTAGCCTCCATCACCCTGACATCATC-- 579
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 580 -----GTTACTGAACATGCTAATCAGGCCCAAGGAG 609
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 610 ACTCTGTATGAGATTACATAGACAATAACGCGCATCGTCTGTCTGTCGGCGGAGATGGT 669
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 670 ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGCGGCTCGAC 729
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 730 CAGAACACACCCCGGGCTGTGCTGCCAGTAGCTCCGGATTGGAATCAITCCCGCA 789
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 790 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACACGACGACGAGAACTCTCGCG 849
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 850 CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACAGC 909
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 910 ACCTCTCTCGCTACTCCGTGCTCCCTGTGGCTTACGGCTTCTACGGGGACATCATCAAG 969
Qy 181 AspSerGluLysLysArgTyrLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 970 GACAGTCAGAAGAACCGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 1029
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 1030 TTCCTCTCCACCACTGCTATGAAGGGACAGTGTCTTCTCTCTCTGACACACACCGGTG 1089
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 1090 GGATCTCCAAGGGATAGGAAGCCCTGCCGGCAGGAGTCTTTGTTTTCAGGCAAGCAAG 1149
Qy 241 GlnGlnLeuGluGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 1150 CAGCAGCTGGAGGAGGAGCAGAGAAGCACTGTATGGTTTGGAGCTGCGGAGGACGTG 1209
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAlaThrAsnMetSerCys 280
```


Db 1210 GAGGAGTGGCAAGTCGTCTGTGGAGAGTTTCTGGCCATCAATGCCACAACATGTCCTGT 1269
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1270 GCTTGTGCGCGAGGCCAGGCCCTCTCCCGGCTGCCCACTTGGGAGCGGTCTTCT 1329
QY 301 AspLeuLeuLeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuLeuHis 320
Db 1330 GACCTCATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCAC 1389
QY 321 ThrAsnGlnGlnAspGln 326
Db 1390 ACCAACCAGCAGGACCAG 1407

RESULT 8

US-09-969-896-16
; Sequence 16, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 4413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-16

Alignment Scores:

Pred. No.: 1.07e-200 Length: 4413
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 95.54% Indels: 11
DB: 10 Gaps: 1

US-10-631-958-2 (1-326) x US-09-969-896-16 (1-4413)

QY 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 463 CCNAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGCAAGGCGGATA 522
QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 523 TATGAAGAAGAAAGTGGCACCAGCTGTCACCTTAGCTTCATCACCAGCTGACATCATC 579
QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 580 -----GTTACTGAACATGCTAAATCAGGCCCAAGGAG 609
QY 61 ThrLeuTyrGluLeuAsnIleAspLysTyrAspGlyIleValCysValGlyLysAspGly 80
Db 610 ACTGTGTATGAGATTAAACATAGACAAATACGACGCGCATCGTCTGTGCGGCGAGATGGT 669
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 670 ATGTTTCAGGAGGTCTGCAGGCTGTGATTGGAGGAGCGCAGGAGCGCGGGTCCAC 729
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 730 CAGAACCACCCCGGGCTGTGCTGCTGCCAGTAGCTCCCGATTGGAATCATTCGCCGA 789
QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140

Db 790 GGGTCAACGGAGCTGGTGTGTATTCTCCACCGTGGCCACGAGGCGAGAAACCTCGGGG 849
QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 850 CTGCATATCGTTGTGGGAGCTCGCTGGCCATGGATGTGTCTCAGTCCACCACAACAGC 909
QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 910 ACATCTCTTCCTACTCTCGTGTCCCTGTGGGCTACGGCTTCTACGGGGGACATCATCAAG 969
QY 181 AspSerGluLysLysArgTyrTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 970 GACAGTGAGAAGAAACGGTGGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 1029
QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 1030 TTCTCTCTCCACCACTGCTATGAAGGAGACAGTCTCTCTCCCTGCGACAACACACAGGTG 1089
QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 1090 GGATCTCCAAAGGATAGGAAGCCCTGCCGGCGCAGGATGCTTTGTTTGAGGCAAGCAAG 1149
QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 1150 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTGCCGAGACGTG 1209
QY 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 1210 GAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAACATGTCTGT 1269
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1270 GCTTGTGCGCGAGGCCCGCCAGCGGCTCTCTCCCGGCTGCCACCTTGGGAGACGGGTCTTCT 1329
QY 301 AspLeuLeuLeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuLeuHis 320
Db 1330 GACCTCATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCAC 1389

RESULT 9

US-10-631-958-16
; Sequence 16, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 4413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-16

Alignment Scores:

Pred. No.: 1.07e-200 Length: 4413
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 95.54% Indels: 11

```
DB: 19 Gaps: 1
US-10-631-958-2 (1-326) x US-10-631-958-16 (1-4413)
QY 1 ProlysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
DB 463 CCAAAGCATTTACTGGTATTTATCAACCGTTTGGAGGAAAAGGACAAGGCAACGGGATA 522
QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
DB 523 TATGAAAGAAAGTGGCACCACCTGTTACCTTAGCTCCATCACCACATCATCATC--- 579
QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
DB 580 -----GTTACTGAACATGCTAATCAGGCCCAAGGAG 609
QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
DB 610 ACTCTGTATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTGTGGCGGAGATGGT 669
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
DB 670 ATGTTACAGAGGTGCTCGACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGTGCAC 729
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
DB 730 CAGAACCAACCCCGGGCTGTGCTGTCGCCAGTAGTACCTCCGGATTGGAATCATTTCCGCA 789
QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
DB 790 GGGTCAACGGGATCGCTGTGTACTCCACCGTGGGACACGACGACGAGAAACCTCGCGG 849
QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
DB 850 CTGCATATCGTTTGGGACTCGCTGGCCATGATGATGCTCTCAGTCCACCAACACAGC 909
QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
DB 910 ACACCTCTTCGCTACTCCGCTGTCCCTGTGGCTACGGCTTCTACGGGGACATCATCAAG 969
QY 181 AspSerGluLysLysArgTyrPheGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
DB 970 GACAGTGAGAAAGAACCGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 1029
QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
DB 1030 TTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTCCCTGCACACACACAGGTG 1089
QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
DB 1090 GGATCTCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGTCAGGCCAAGCAAG 1149
QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
DB 1150 CAGCAGCTGGAGGAGGAGCAGAAAGCACTGTATGGTTTGGAGCTGCGGAGGACGCTG 1209
QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
DB 1210 GAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACCAACATGTCCTGT 1269
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
DB 1270 GCTTGTCCCGAGAGCCCGAGGGCCCTCTCCCGGGCTGCCACTTGGGAGACGGGTCTTCT 1329
QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
DB 1330 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1389
QY 321 ThrAsnGlnGlnAspGln 326
DB 1390 ACCAACCAAGCAGGACCAG 1407
QY 1 ProlysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
DB 479 CCAAAGCATTTACTGGTATTTATCAACCGTTTGGAGGAAAAGGACAAGGCAACGGGATA 538
QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
DB 539 TATGAAAGAAAGTGGCACCACCTGTTCACTTACCTTACCTCCATCACCACATCATCATC--- 595
QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
DB 596 -----GTTACTGAACATGCTAATCAGGCCCAAGGAG 625
QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
DB 626 ACTCTGTATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTGTGGCGGAGATGGT 685
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
DB 686 ATGTTTACGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGGTGCAC 745
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
DB 746 CAGAACCAACCCCGGGCTGTGCTGGTCCCACTAGTACCTCCGATTGGAATCATTTCCGCA 805
QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
DB 806 GGGTCAACGGGACTCGCGTGTGTTACTCCACCGTGGGACCGACGACGAGAAACCTCGCGG 865
QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
DB 866 CTGCATATCGTTGTGTGGGACATCGCTGGCCATGATGCTCTCAGTCCACCAACACAGC 925
QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
DB 926 ACATCTTCTCTACTCCGTTGCTCCCTGGGCTACGGCTTCTACGGGGACATCATCAAG 985
QY 181 AspSerGluLysLysArgTyrPheLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
DB 986 GACAGTGAGAAAGAACCGTGTGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGACC 1045
QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
```

Db 1046 TTCTCTCCCACTGCTATGAAGGACAGTGCTCTTCTCCCTGCACACACACGGTG 1105
QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 1106 GGATCTCCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTGCAAGCAAGCAAG 1165
QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 1166 CAGCAGCTGGAGAGAGAGCAGAAAGACACTGTATGTTTGGAAAGCTGCCGAGGACGTG 1225
QY 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 1226 GAGGAGTGCGAAGTCTGTCTGGAGATTCTTGGCCATCAATGCCACAAACATGTCCTGT 1285
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1286 GCTTGTGCGCGAGAGCCCGAGGGCCCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCT 1345
QY 301 AspLeuIleLeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1346 GACCTCATCTCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1405
QY 321 ThrAsnGlnGlnAspGln 326
Db 1406 ACCACACAGCAGGACCAG 1423

RESULT 11

US-10-120-988-148
; Sequence 148, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt FL_genes Version 2.0
; SEQ ID NO 148
; LENGTH: 4432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129)..(2817)
US-10-120-988-148

Alignment Scores:

Pred. No.: 8,686-200 Length: 4432
Score: 1633.50 Matches: 313
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.01% Mismatches: 11
Query Match: 95.14% Indels: 11
DB: 17 Gaps: 1

US-10-631-958-2 (1-326) x US-10-120-988-148 (1-4432)

QY 1 ProllysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 1591 CCAAAGCATTTACTGGTATTTATCAACCCGTTTGAGGAAAAGGCAAGGCAAGCGGATA 1650
QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 1651 TATGAAGAAAGAGTGGCAACCACTGTTACCTTAGCTCCATCACCACCTGACATCATC--- 1707

QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 1708 -----GTTACTGAACATGCTAATACAGCCCAAGGAG 1737
QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 1738 ACTCTGTATGAGATTACATAGACAAATACAGCGGCATCGTCTGTGTTCGGCGGAGATGGT 1797
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 1798 ATGTTACGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCCGGGTGAC 1857
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 1858 CAGAACCAACCCCGGGCTGCTGTCCTCCCACTAGCTCCGGATTGGAATCATTTCCCGCA 1917
QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 1918 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACACGACGACGACGACAACTCTGGCG 1977
QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 1978 CTGCATATCGTTGTGGGACTCGCTGCCCATGATGTCTCAGTCCACCAACACAGC 2037
QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 2038 ACCTCTCTCGTACTCGGTGCTCCCTGCGGTACGGCTTCTACGGGGACATCATCAAG 2097
QY 181 AspSerGluLysLysArgTTPLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 2098 GACAGTGAGAAAGAAACCGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 2157
QY 201 PheLeuSerHisHisCysTyrTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 2158 TTCTCTCTCCCACTCGTATGGAGGACAGTGTCTTCTCTCCCTGCACAAACACACAGGTG 2217
QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 2218 GGATCTCCAAGGATAGGAAAGCCCTGCCGGCAGGATGCTTTGTTGCGAGCAAGCAAG 2277
QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 2278 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGTTTGGAAAGCTGCCGAGGACGTG 2337
QY 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 2338 GAGGAGTGGAAGTCTGCTGTGGGAAAGTTCTGCGCATCAATGCCACAAACATGTCCTGT 2397
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 2398 GCTTGTGCGCGAGCGCCCGAGGGCTCTCCCGGCTGCCCACTTGGGAGAGCGGTCTTCT 2457
QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 2458 GACCTCATCTCTATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 2517
QY 321 ThrAsnGlnGlnAspGln 326
Db 2518 ACCAACACAGCAGGACCAG 2535
RESULT 12
US-10-315-597A-1
; Sequence 1, Application US/10315597A
; Publication No. US20030162206A1
; GENERAL INFORMATION:
; APPLICANT: Sugiyura, Masako
; APPLICANT: Kono, Keita
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Ceramide
; FILE REFERENCE: 02658CIP/HG
; CURRENT APPLICATION NUMBER: US/10/315,597A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: JP 2000-178039

Db 308 AAGGACAGTGAGAAGAAACGGTGGTTCCTGGCCAGATACGACTTTTCAGGTTTAAAG 367
QY 200 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 219
Db 368 ACCTTCCTCCACCACCTGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACG 427
QY 220 ValGlySerProArgAspArgLysProCysAisqAlaGlyCysPhe 234
Db 428 GTGGGATCTCCAAGGATAGGAAGCCCTGCCGGGCAAGATGCTTT 472

RESULT 14

US-10-631-958-4
; Sequence 4, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Koestda, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-4

Alignment Scores:
Pred. No.: 1,07e-94 Length: 474
Score: 811.00 Matches: 154
Percent Similarity: 99.35% Conservative: 0
Best Local Similarity: 99.35% Mismatches: 1
Query Match: 47.23% Indels: 0
DB: 19 Gaps: 0

US-10-631-958-2 (1-326) x US-10-631-958-4 (1-474)

QY 80 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 99
Db 8 GGTATGTTCCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGCAGAGGAGCGCGGGTC 67
QY 100 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro 119
Db 68 GACCAGAACCCCGCGGCTGTGCTGGTCCCAGTAGCTCCGGATTGGAAATCATTCCTCC 127
QY 120 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 139
Db 128 GCAGGGTCAACGACTGCGTGTGTACTCCACCGTGGGACCAGGACGCAGAAACCTCG 187
QY 140 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHisAsn 159
Db 188 GCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAAC 247
QY 160 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 179
Db 248 AGCACACTCTTCGCTACTCCGTGTCCTGTGCTGGGTACGGCTTCTACGGGGACATCATC 307
QY 180 LysAspSerGluLysIleArgTrpLeuGlyLeuAlaAaTyrAspPheSerGlyLeuLys 199
Db 308 AAGGACAGTGAGAAGAAACGGTGGTCTGTCCAGATACGACTTTTCAGGTTTAAAG 367
QY 200 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 219
Db 368 ACCTTCCTCCACCACCTGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACG 427

QY 220 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 234
Db 428 GTGGGATCTCCAAGGATAGGAAGCCCTGCCGGGCAAGATGCTTT 472

RESULT 15

US-10-425-115-176413
; Sequence 176413, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 176413
; LENGTH: 2657
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MMT4577_92481C.1
US-10-425-115-176413

Alignment Scores:
Pred. No.: 5.6e-57 Length: 2657
Score: 529.00 Matches: 131
Percent Similarity: 45.97% Conservative: 57
Best Local Similarity: 32.03% Mismatches: 107
Query Match: 30.81% Indels: 114
DB: 20 Gaps: 11

US-10-631-958-2 (1-326) x US-10-425-115-176413 (1-2657)

QY 1 ProlHisLeuLeuValPheIleAsnProPheGlyLysGlyGlnGlyLysArgIle 20
Db 556 CCAAAGAACTTAATGGTATTTGTTTCATCCATTGTGTGGAAAGGTAGAGGGTCAAGTAAC 615
QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 616 TGGGAA---ACAGTGTATCCGTTATTTCTGTAGACGAAGATAAACAAGGTGATA--- 669
QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 670 -----GTGACAGAGAGAGCAGGCGCATGCATATGCAC 699
QY 61 Thr-----LeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 77
Db 700 ACACATCATCATTCATAGATATAGAACTGAAGGCATTTGATGGTGTGTTGTCAGTGGGT 759
QY 78 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArg----- 95
Db 760 GGCACGGTCTATTTAATGAAATTCCTGAATGGGACTACTAGTTCTAGGCACAAGGTTTCA 819
QY 95 ----- 95
Db 820 TATCCCCCACTCCTGAGGGGTTTGGATATGTTGGAACAGCTGATACTGTGAAGAACAG 879
QY 95 ----- 95
Db 880 ACAAGAAATTGGGTAACTTTTAGTAAACCCACACACAGATTCCTGGGAATGCGGTGTTCTA 939
QY 96 -----SerAlaGlyValAspGln-----AsnHisProArgAla 106
Db 940 GGGAAATCCAAGTAATGTGATGACCAAGAACCTCTCTTTTCACTTCACATCCTAGCCAG 999
QY 107 ValLeuValProSerSer----- 112
Db 1000 CTAGAGATACCATCAAAATTCCTACTAAAGAACCACTGATTGGAGATCAGGATAACCCAGTT 1059

```
Qy 113 -----LeuArgIleGlyIleIleProAlaGlySerThrAspCysVal 126
Db 1060 TCCTTCCCAATGACTGGTTTAGGCTTGGCATAAATCTCTCTGGCTCAACTGATGCTATT 1119
Qy 127 CysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGly 146
Db 1120 GTTCTCAGCAACAACCTGGGGAGAGATCCTGTCACCTCTGCTCTGCTTATTATTCTTGGC 1179
Qy 147 AspSerLeuAlaMetAspValSerSerValHisIleAsnSerThr----- 161
Db 1180 AGAAAGGTGTCACTTGTATATAGCTCAAGTCCTTAGTGGAAACTAGCCAGCAGCTGAG 1239
Qy 162 -----LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 178
Db 1240 GTTTTGCCTACTGTACGCTATGCTGCTTCATTTCAGGTTATGGCTTTTATGGAGAAGTT 1299
Qy 179 IleIysAspSerGluLysGlyArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 198
Db 1300 ATCAGGGAGAGCGAAAGTACCGGTGGATGGGTCTGCACGTTATGACTTTTCTGGAACA 1359
Qy 199 LysThrPheLeuSerHisCysTyrGlyThrValSerPheLeuProAlaGlnHis 218
Db 1360 ATGGTCTTCTGAAGCAGGTCAATATGGGCAAGGTTGCTTTCTCGATACTCCATAC 1419
Qy 219 Thr-----ValGlySerProArgAsp----- 225
Db 1420 ACTCATTCGCTTACAGCATCAGCACAGGATGATATACTGGAGCACCACTACAATCT 1479
Qy 226 -----ArgLysPro-----CysArgAlaGlyCysPheValCysArgGlnSer 239
Db 1480 CGTTGGAAGAAACCTCGCAAAATAATCTGCAGAACAAATTGCTTTGTATGCAAGAGGCT 1539
Qy 240 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 259
Db 1540 TCAACATCTGGACAAAAATCCAGATGATGTTGCAGACAACTCCAGAACAAATATCGAGAAC 1599
Qy 260 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 279
Db 1600 ---CAGAAAGTGGTCTGGTCCGAGGACATTTCTCAGCGTCGGTGCAGCTGTCTATTCA 1656
Qy 280 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 299
Db 1657 TGTCTAAACGAAAGAGCCCTGATGGCTGTAGTGTGACGACACACCTTTCGGATGGTTT 1716
Qy 300 SerAspLeuIleLeuIleArgLysCys 308
Db 1717 CTTTCATCTCTTGTCTAAGAGACTGC 1743
```

Search completed: September 6, 2005, 10:47:27
Job time : 951.291 secs

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	486.5	28.3	549	2	T33517	hypothetical prote	
2	365.5	21.3	1240	2	T05162	hypothetical prote	
3	226	13.2	458	2	T38776	hypothetical prote	
4	220.5	12.8	687	2	S51398	hypothetical prote	
5	219.5	12.8	473	2	T19707	hypothetical prote	
6	185.5	10.8	310	2	AG1665	hypothetical prote	
7	181.5	10.6	624	2	S67059	hypothetical prote	
8	178.5	10.4	310	2	AI1293	hypothetical prote	
9	131	7.6	306	2	AH1769	conserved hypothet	
10	130	7.6	309	2	AH1528	conserved hypothet	
11	129	7.5	303	2	F69795	conserved hypothet	
12	125.5	7.3	295	2	A83894	hypothetical prote	
13	125	7.3	294	2	G95120	conserved hypothet	
14	123.5	7.2	309	2	AE1171	conserved hypothet	
15	121	7.0	306	2	AE1394	conserved hypothet	
16	121	7.0	311	2	C97990	conserved hypothet	
17	116.5	6.8	364	2	F84898	hypothetical prote	
18	113	6.6	295	2	D83734	hypothetical prote	
19	111.5	6.5	345	2	E69678	involved in polyke	
20	110	6.4	433	2	S75948	hypothetical prote	
21	107	6.2	1028	2	A96719	hypothetical prote	
22	105	6.1	315	2	A89978	conserved hypothet	
23	105	6.1	732	2	T16422	hypothetical prote	
24	102	5.9	343	2	JC7183	cathespin Q [EC 3.	
25	101	5.9	309	2	H70961	hypothetical prote	
26	97.5	5.7	309	2	H69995	hypothetical prote	
27	95	5.5	455	1	B25493	indoleacetamide hy	
28	93	5.4	650	1	J14C50	fibroblast growth	
29	92.5	5.4	304	2	F72386	conserved hypothet	

A;Cross-references: UNIPROT:Q92A05; GB:AL592022; PIDN:CAC97095.1; PID:g16414366; GSPDB:8
A;Experimental source: strain Clip11262

C;Genetics:
A;Gene: lin1865

Query Match 10.8%; Score 185.5; DB 2; Length 310;
Best Local Similarity 23.0%; Pred. No. 1.6e-08;
Matches 73; Conservative 46; Mismatches 114; Indels 85; Gaps 11;

Qy 2 KHLVFNPFPGKGOGKRIYERKVAPLFTLASITTDIIGNKFVYNVVEVITEHANQAKET 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3 KHARVIYNP-----TSGREIIKKNLADVLS-----ILEQAGVYTSAAHTTAEPGDACHA 51
Qy 62 LYENIDKYDGI VCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAG 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
52 AEEAVRDYDLVAAGGGDTINEVINGIAEQPYRP-----KVGIIPTG 94
Qy 122 STDCCVYSTVTGSDAETSALHIIVVGDSLAMDYSSVVHHNSTLLRYSVSLLGVGYFDIKD 181
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
95 TTDFARAHVPRDIVIKATKI IAGOSVANDLGKA--NDT---YFINIGGGRLUTELTYD 149
Qy 182 --SBKKRWLGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGFCVCQRQS 239
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
150 VPSRLKTMQLAYLVKG-----IEMLPS-----LKAT 177
Qy 240 KQLEEBEOKKALYLEAEADVEEMVQVKGFPLAINATNMSCACRRSPRGSLSPAHLGDGS 299
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
178 KVKEYDQ-----GVPEGEVM---FFLLGLTNISGGFEK----IAPDAKLDDGK 219
Qy 300 SDLILIRKCSRNFNLRFL 317
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
220 FSLIIVKKNLAEFIRLV 237

RESULT 7

S67059
hypothetical protein YORI171c - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein O3615
C;Species: *Saccharomyces cerevisiae*
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C;Accession: S67059
R;Bordome, R.; Camaseses, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarasov, I.A.; Winsor
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67032
A;Accession: S67059
A;Molecule type: DNA
A;Residues: 1-624 <BOR>
A;Cross-references: UNIPROT:Q12446; EMBL:Z75078; NID:g1420415; PID:g1420417; GSPDB:GN000
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:LCB4; MIPS:YORI171c
A;Cross-references: SGD:S0005697
A;Map position: 15R

Query Match 10.6%; Score 181.5; DB 2; Length 624;
Best Local Similarity 25.2%; Pred. No. 9.1e-08;
Matches 62; Conservative 43; Mismatches 96; Indels 45; Gaps 7;

Qy 2 KHLVFNPFPGKGOGKRIYERKVAPLFTLASITTDIIGNKFVYNVVEVITEHANQAKET 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
227 RSILVINPHGGKGTAKNFLFTKARPILVES-----GCKIRIAY----TKVARHAIDI 275
Qy 62 LYENIDKYDGI VCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAG 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
276 AKLDJDSKYDTIACASGGDI PVVEINGLYRRPDR---VDAFNKLAIVTQ-----LPCG 324
Qy 122 STDCCVYSTVTGSDAETSALHIIVVGDSLAMDY---SSVHHNSTLLRYSVSLLGVGYFGYDI 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
325 SGNAWSISCHWNNPNSYAALCLVKSIEFTIDL MCCSQPSMYNEWRPLSFLSQYGVTAES 384
Qy 179 IKOSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGFCVCQRQ 238
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
385 DINTEFFRMWGPPVFNLG-----VAPNIIOG-----KKYCEVFVKYAAK 424

[illegible]

C;Accession: AB3894
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai,
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: AB3650; MUID:20512582; PMID:11058132
A;Accession: AB3894
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-295 <STO>
A;Cross-references: UNIPROT:Q9KBH4; GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BAB056
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1953

Query Match 7.3%; Score 125.5; DB 2; Length 295;
Best Local Similarity 21.2%; Pred. No. 0.0027;
Matches 68; Conservative 46; Mismatches 108; Indels 99; Gaps 16;

Qy 6 VFINPFGGKGGKRIYERKVAPLFTLASITTDIIGNKFVYNYVEVITEHANQAKEYIYEI 65
: : | : | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 LIVNKASGNGKGQRWKK-----VEVELQIRNTPYLVRP---TSGSGHATTIVKEL 51

Qy 66 NIDKYDGVICVCGDMFSEVLHLIGRTORSAGVDONHPRAVLPVPSLRGIIPAGSTDC 125
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 52 LTEGVKTIIAIVGGDTINEAVGLV-----NH-----RVV----LGIIIPAGS--- 89

Qy 126 VCYSTVGTSDAE-----TSALHVVGDSLAMDVSSVHHNSTLLRYSVSLLYGFGVD 177
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 90 -----GNDFARCLNIPMHYEKALHRIP-ENKQKKVDLHLGQ---RHCLTVTGIGFDGK 139

Qy 178 IIK---DSEKKRWLGLARYDFSGLKTFILSHHCYEGTVSFPLAQHTVGSPPDRKPCRAGCF 234
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 140 IAKTWNEAIYNWFN--QFGFGLSVWLS-----MLEVLKDYRP----- 176

Qy 235 VCRSKQOLEEQKALYGLEAAEDVEEVQVCGFKFLAINATNMSCACRRSPRLGPSAAH 294
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 177 ----TNIQITVDGKELFFS-----GWLVLA-----VANSPNYGGGIR-----ICPEAS 215

Qy 295 LGDGSSDLILIRKCSRNFRLR 315
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 216 YDDGLLNLCVVHGMSKWQLLR 236

RESULT 13
G95120
conserved hypothetical protein TIGR00147 SP1045 [imported] - Streptococcus pneumoniae (s)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: G95120
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a Virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: G95120
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-294 <EUR>
A;Cross-references: UNIPROT:Q97QZ6; GB:AE005672; PIDN:AAK75160.1; PID:g14972520; GSPDB:G
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SPI045

Query Match 7.3%; Score 125; DB 2; Length 294;
Best Local Similarity 23.4%; Pred. No. 0.003;
Matches 49; Conservative 37; Mismatches 85; Indels 38; Gaps 8;

Qy 2 KHLVFNPFCKGGKGRKIYERKVAPLFTLASITTDIIGNKP---YNNYVEV-ITEHANQA 58
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 2 KKAMVIINPTSG-GEKALDYKEK-----LENKAKEYFEYVETKITEKALDA 46

RESULT 12
A83894
hypothetical protein BH1953 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence revision 01-Dec-2000 #text change 09-Jul-2004

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 6, 2005, 07:01:07 ; Search time 3134.86 Seconds
(without alignments)
3958.373 Million cell updates/sec

Title: US-10-631-958-2
Perfect score: 1717
Sequence: 1 PKHLLVFNPFGCGKQGRRI.....KCSRNFNLFRLHRTNQDQ 326

Scoring table: BLOSSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10631958/runat_02092005_165816_4562/app.query.fasta_1.1941
-DB=BST -OPMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blossum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CGN 1 1 9235 @runat_02092005_165816_4562 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1500.5	87.4	1824	3 AK042077	Mus muscu
2	1500.5	87.4	4248	3 AK052269	Mus muscu
3	1455.5	84.8	1063	4 BM479389	AGENCOURT
4	1350.5	78.7	1059	5 BQ054406	AGENCOURT
5	1283	74.7	797	7 CK000755	AGENCOURT
6	1267.5	73.8	1047	5 BQ057191	AGENCOURT
7	1176	68.5	758	6 CB246749	UI-M-F10-
8	1073.5	62.5	1078	5 BQ063738	AGENCOURT
9	1041	60.6	820	6 CD655311	AGENCOURT

10	1012	58.9	581	5	BP224560	BP224560
11	980.5	57.1	584	5	BP10011	BP10011
12	965.5	56.2	584	5	BP30990	BP30990
13	951.5	55.4	573	7	CF138275	UI-HF-BNO
14	942	54.9	541	7	CN296311	170005321
15	925	53.9	725	5	BQ214294	603756937
16	917	53.4	550	7	CF138634	UI-HF-BNO
17	916	53.3	576	6	CB608292	AGNNNUC:N
18	905.5	52.6	779	5	BQ265377	603374309
19	898	52.3	713	7	CK000525	AGENCOURT
20	892	52.0	661	7	CN296312	170005321
21	873	50.8	674	1	AJ739034	AJ739034
22	861.5	50.2	553	6	CA578972	K0731A08-
23	831	48.4	545	6	CB611947	AMGNNUC:N
24	828	48.2	488	5	BQ430459	UI-HF-BNO
25	811	47.2	474	2	AW503999	UI-HF-BNO
26	809	47.1	902	5	BQ373554	603811294
27	795	46.3	653	5	BX952302	DKF20781L
28	761	44.3	759	5	BQ290807	604162346
29	749	43.6	759	5	BQ264162	603814452
30	741	43.2	732	7	CF135528	UI-HF-BNO
31	705.5	41.1	735	7	CF535441	UI-M-GHO-
32	695.5	40.5	467	4	BM484519	538413 MA
33	694.5	40.4	638	1	AJ739022	AJ739022
34	688.5	40.1	452	6	CA579952	K0744G05-
35	640	37.3	819	7	CN235868	WLB086C10
36	637.5	37.1	876	7	CN235868	WLB086C10
37	627	36.5	567	1	AJ739023	AJ739023
38	591	34.4	359	5	BQ428926	UI-HF-BNO
39	588	34.2	763	7	CK364327	AGENCOURT
40	587.5	34.2	725	5	BQ225411	603399474
41	570	33.2	607	6	CB581157	AMGNNUC:N
42	570	33.2	656	4	BQ673686	BJ673686
43	555	32.3	660	4	BJ620045	BJ620045
44	508.5	29.6	329	1	AA355581	EST64087
45	500.5	29.1	310	6	CD631566	56059545H

ALIGNMENTS

RESULT 1	AK042077	AK042077	1824 bp	linear	HTC 03-APR-2004
LOCUS	AK042077	Mus musculus	3 days neonate thymus cDNA	RIKEN full-length enriched library, clone:A630056D11 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence.	
DEFINITION	AK042077.1	GI:26334912	HTC; CAP trapper.		
ACCESSION	AK042077	GI:26334912	Mus musculus (house mouse)		
VERSION	AK042077.1	GI:26334912	Mus musculus		
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1	Carninci,P. and Hayashizaki,Y.			
AUTHORS	1	Carninci,P. and Hayashizaki,Y.			
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.			
AUTHORS	2	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,			

Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE

JOURNAL

REFERENCE

AUTHORS

5 Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1824)

JOURNAL

REFERENCE

AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saichou,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp].

TITLE

JOURNAL

REFERENCE

AUTHORS

URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/.

FEATURES

source

1..1824
/location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:A630056D11"
/db_xref="taxon:10090"
/clone="A630056D11"
/tissue_type="thymus"
/clone_libs="RIKEN full-length enriched mouse cDNA library"
/dev_stage="3 days neonate"
85..1680
/note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTF|Q9UG55, evidence: FASTA, 86.5%ID, 100%length, match=945) putative"
/codon_start=1
/protein_id="BAC31157.1"
/db_xref="GI:26334913"
/translation="MGAMGAAEPLHSLVLMVKRRRCVSLSEPARALLRWRSPEPGPSA

CDS

PGADARSLVLSIEIIAIVEEKDDCEKHASSGRWHKVENPFAFVHRVKKRVHHRWKWARV
TFSADEQLCHLWQLTLLGLESITSRPKLLVFINPFGGQCKRIYETKTVAPLFTL
ASTTTEII TEHANOAKETIYEINTSDYIGVCGDGMFSEVLHGVIKRTQOSAGID
PNHPRALVPSLRIIGIPAGSDCYCVSTVGTNDAETSAHIIIGDSLADYSSVHV
HNTLLRYSVLLGYPGYDLIKDSEKRWMLVRYDFSGLTKFTLSHOYVGTGKFLPA
QHTVGSPRNDKCRAGFCVCRSQKQLEBEKALYGLNEAEEMEEQVVTGKFLPA
ATNMSCACPRSPGSLSPFAHLGGSDLLILRKCSRPNFLRLRIHTNQDQDFTFV
EYVRVKKFHTSKHVEDENDSKEQBKQKFKGI CKDRPSCSASRSSMNCNDCGVRMHS
PAIEVRVHCQLVRLFARGIEES"

ORIGIN

Alignment Scores:

Pred. No.: 1,04e-154 Length: 1824
Score: 1500.50 Matches: 282
Percent Similarity: 92.64% Conservative: 20
Best Local Similarity: 86.50% Mismatches: 13
Query Match: 87.39% Indels: 11
DB: 3 Gaps: 1

US-10-631-958-2 (1-326) x AK042077 (1-1824)

Qy 1 ProlyserHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIle 20

Db 472 CCGAAGCAGCTTGGTGGTATTTCATCAACCCCTTCGAGGGAAAGTCAAGGCAACGCCATC 531

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40

Db 532 TATGAAAAACAGTGGCGCCTCTGTTTACCTTGGCTCCATCTACTACGGAGATCATC-- 588

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60

Db 589 -----ATTACAGCATGCCAACCAAGCAAGGAG 618

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80

Db 619 ACTTTATACGAGATCAACACAGACAGCTATGATGGCATCGTGTGCTAGGTGGGAGCGGC 678

Qy 81 MetPheSerGluValLeuHisGlyIleGlyArgThrGlnArgSerAlaGlyValAsp 100

Db 679 ATGTTACGCGAGGTGCTGCGTGGTATTGGGAGGACGACGACGCGCTGGTATCGAC 738

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120

Db 739 CCCAATCACCCCGAGCGGTGCTGGTCCCGCATCCCTCAGGATCGGCATCATACCCGCA 798

Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140

Db 799 GGGTCCACAGATTGTGTGTACTCAACAGTGGGCACAAACGACGACGACATCGGCT 858

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160

Db 859 TTGCACATCATTTATGGGACTCCTCGCAATAGACGTCCTCTGTGCACTACCAATAAC 918

Qy 161 ThrLeuLeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180

Db 919 ACGTGTGCGGTACTCGGTTTCTCTGCTGGGCTACCGTTTCTACGGGGAGATTAAACAAG 978

Qy 181 AspSerGluLysLysArgTyrLeuAlaArgTyrAspPheSerGlyLeuLysThr 200

Db 979 GACAGTGAAAGAAACGGTGGATGGGCTCTGTCGGTATGATTTCTCAGGGGTTGAAGACC 1038

Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220

Db 1039 TTTCTCTCTCATCATGACTATGAAGGGACACTGCTCTCTCCAGCACAGCACACCGGTG 1098

Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240

Db 1099 GGATCTCCACGGGCAATAAACCCCTGCGGGCTGGTCTCGTGTGTCAGGACGACGAGCAAG 1158

Qy 241 GlnGlnLeuGluGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260

Db 1159 CAACAGCTGGGAAGAAGAGAGAAAGCCCTGTATGGCTTGGAGAACGCCGACGAGGAATG 1218


```
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 764 -----ATTACAGAGCATGCCAACACAGCCAGGAG 793

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyAspGly 80
Db 794 ACTTTATACGAGATCAACACAGACAGCTATGATGCGCATCGTGTGCTAGGTGGGACGC 853

Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 854 ATGTTACGCGAGGTCTCATCGGGGTGATGGGAGGACGACGAGCGCTGGTATCGAC 913

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db 914 CCCATACCCCCGAGCGCTGCTGGTGCCTAGTACCTCAGATCGGATCATACCCGCA 973

Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 974 GGGTCCACAGATTGTGTGTTACTCAACAGTGGGCACAAACGACGACAGACATCGCT 1033

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSer 160
Db 1034 TTGCACATCATATTGGGGACTCACTGGCAATAGACGTCTCTGTGCATCACCATAAC 1093

Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuLys 180
Db 1094 ACGTGTCTCGGTACTCGGTTTCTCTGTGGGCTACGGTTTCTACGGGGACTTATCAAG 1153

Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 1154 GACAGTGAAGAAAGAACGGTGGTGGCTCGCGGTATGATTTCTCAGGGTTGAAGACC 1213

Qy 201 PheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 1214 TTTCTCTCTCATCAGTACTATGAAGGGACACTGTCTCTCCAGCAGCACACACGGTG 1273

Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 1274 GGAITCCACGGGACAATAAACCTTCGCGGCTGGTCTCTGTGTGAGCAGACGACGAG 1333

Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 1334 CAACAGCTGGAAGAAGAAGAAAGCCCTGTATGGCTGGAGAACCCGCGAGGAATG 1393

Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 1394 GAAGAGTGGCAAGTGACATGTGGGAAGTTCTTGCGCCATCAATGCCAACATGTCTCTGT 1453

Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1454 GCTTGTCTCGGAGCCCTCGGGGCTGTCCCATTTGGCCATCTGGGAGATGGGTCTTCT 1513

Qy 301 AspLeuIleLeuAlaArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1514 GACCTCATCTTATCCGGAAGTCTCCAGGTTCAACTTCTCTGAGATTCTCATCCGGCAC 1573

Qy 321 ThrAsnGlnGlnAspGln 326
Db 1574 ACGAACACGAGGACGAC 1591
```

```
RESULT 3
BM479389
LOCUS BM479389 1063 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6418742 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5502281
5', mRNA sequence.
ACCESSION BM479389
VERSION BM479389.1 GI:18528431
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 1063)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12140 row: c column: 18
High quality sequence stop: 665.
FEATURES
source
1..1063
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5502281"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 4,47e-150 Length: 1063
Score: 1455.50 Matches: 283
Percent Similarity: 94.68% Conservative: 2
Best Local Similarity: 94.02% Mismatches: 4
Query Match: 84.77% Indels: 12
DB: Gaps: 1

US-10-631-958-2 (1-326) x BM479389 (1-1063)

Qy 26 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGlyAsnLysPheTyrVal 45
Db 1 GCACCACTGTTCACCTTAGCCTCCATCACCACCTGACATCATC----- 42

Qy 46 AsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIle 65
Db 43 -----GTTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATT 87

Qy 66 AsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSerGluVal 85
Db 88 AACATAGACAAATACGACGCGCATCGTCTGTGTGCGCGGAGATGTTATGTTCCAGCGAGGTG 147

Qy 86 LeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArg 105
Db 148 CTGCACGCTCTGATTGGGAGACGACAGAGGAGCGCGGGGTCCAGCAACACACCCCGG 207

Qy 106 AlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCys 125
Db 208 GCTGTCTGTGTCCTCCAGTAGCTCCGATTGGAAATCAATCCCGCAGGGTCAACGGACTGC 267

Qy 126 ValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValVal 145
Db 268 GTGTGTTACTCCACCGTGGGCACCGACGACGAGAAACCTCGGGCTGCATATCGTTGTT 327

Qy 146 GlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyr 165
Db 328 GGGGACTCGCTGGCCATGGATGTGCTCTCAGTCACCCACACACACACTCTCTTCGGCTAC 387

Qy 166 SerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLys 185
Db 388 TCCGTGTCCCTGTCTGGGTACGGCTTCTACGGGACCATCATCAGGACGACAGTGAAGAAA 447

Qy 186 ArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHis 205
```



```

Db      448  CGGTGGTGGGCTTGGCAGATACGACTTTTCAAGTTTAAAGACCTTCTCTCCACCAC 507
Qy      206  CysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAep 225
Db      508  TGTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACCGTGGGATCTCAAGGGAT 567
Qy      226  ArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGlu 245
Db      568  AGAAGCCCTGCCGGCAGATGCTTTGTTGAGCAAGCAAGCAGCAGCTGGAGGAG 627
Qy      246  GluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluLutrpGlnVal 265
Db      628  GAGCAGAAGAAGCACTGTATGGTTTGAAGCTGCGGAGCAGCTGGAGGATGGCAAGTC 687
Qy      266  ValCysGlyLysPheLeuAlaLeuAlaLeuAlaThrAsnMetSerCysAlaCysArgSer 285
Db      688  GTCTGTGGGAAGTTTCTGCNCATCATGCCACAAACATGCTCTGTCTGTCTGCTGGC 747
Qy      286  ProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeuLeu 305
Db      748  CCCANGGGGCTCTCNCGCCCTGCCACCTTNGAGACGGGTCTTCTGACCTCATCTCATC 807
Qy      306  ArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuArgHisThrAsn-GlnGlnAs 325
Db      808  CGGAATGCTCCAGGNYCAATTTCTGAGATTTCTCATCAGGCACACCAACCACAGAGA 867
Qy      325  p 325
Db      868  C 868

```

```

RESULT 4
BQ054406
LOCUS   AGENCOURT_6771237 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5803668
5', mRNA sequence.
ACCESSION BQ054406
VERSION   BQ054406.1 GI:19813746
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1059)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Lou Staudt
          cDNA Library Preparation: Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LNCM2040 row: a column: 13
          High quality sequence stop: 601.
          Location/Qualifiers

```

```

FEATURES
         1..1059
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5803668"
            /tissue_type="lymphoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_99"
            /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G). Size-selected >500bp for average insert size
            1.8kb. Library constructed by Ling Hong in the laboratory
            of Gerald M. Rubin (University of California, Berkeley)"

```

using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

```

ORIGIN
Alignment Scores:
Pred. No.:      1,9e-138      Length:      1059
Score:          1350.50      Matches:      275
Percent Similarity: 89.49%      Conservative: 6
Best Local Similarity: 87.58%      Mismatches:  14
Query Match:    78.65%      Indels:      19
DB:             5           Gaps:          5

US-10-631-958-2 (1-326) x BQ054406 (1-1059)

Qy      4   LeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArg 23
Db      15  TTACTGGTATTTATCAACCGCTTTGGAGGAAAGGACAAGCAAGCGGATATATGAAAGA 74
Qy      24  LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGlyAsnLysPhe 43
Db      75  AAAGTGGCACCACATGTTTCACTTACCTTCCATCACCACCTGACATCATC----- 122
Qy      44  TyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyr 63
Db      123  -----GTTACTGAACATGCTAATCAGGCCCAAGGAGACTCTGTAT 161
Qy      64  GluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSer 83
Db      162  GAGATTAACTAGACAATAACGACGGCATGCTGTGTGCGGGAGATGGTATGTTTCAGC 221
Qy      84  GluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHis 103
Db      222  GAGGTGCTGCACGCTGCTGATTGGGAGGACGAGAGGCGCGGGGTGACCCAGAACCCAC 281
Qy      104  ProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThr 123
Db      282  CCCCAGGCTGTCTGTGCTCCCGATAGCTCCGGATTGGAATCATTTCCCGAGGGTCAACG 341
Qy      124  AspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIle 143
Db      342  GACTCGGTGTGTTACTCCACCGTGGCCACGACGACGAGCAAGAAACCTCGCGCTGCATATC 401
Qy      144  ValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeu 163
Db      402  GTTGTGGGAGCTCGCTGGCCATGGATGTGCTCCTCAGTCCACCACCAACAGCACACTCCTT 461
Qy      164  ArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGlu 183
Db      462  CGCTACTCCGTGCTCCCTGCTGGGCTTACGGGTCATCATCAAGGACAGTGGAG 521
Qy      184  LysLysArgTTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSer 203
Db      522  AAGAAACCGTGTGGGTCTTGCACAGATACACATTTTCAGGTTTAAAGACCTTCTCTCTCC 581
Qy      204  HisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerPro 223
Db      582  CACCACCTGCTATGAACGAGCAGTGTCTTCTCCCTGCAACACACACGCTGGGATCTCCA 641
Qy      224  ArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeu 243
Db      642  AGGGATAGGAAGCCCTTCCCGGCGAGGATGCTTTGTTTTCAGGGCAAGCAAGCAGCAGCTG 701
Qy      244  GluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluLutrp 263
Db      702  GAGGAGGAGCAGAGAAGCACTGTATGGGTGGAACTGCGGAGGACGCTGGAGGAATGG 761
Qy      264  GlnValValCys-GlyLysPheLeuAlaIleAsnAlaThr-AsnMetSerCysAla---- 282
Db      762  CAAGTCGTCTGTGGGAAGTTCTCTGGCCATTAATGCCCCAACAAACATGCTGCTGCTTGT 821
Qy      282  ysArgArgSerProArgGlyLeuSerProAlaAla-----HisLeuGly---AspGlyS 299
          :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

```

Db      822  CCGCCGGACCTCCAGGGGCGCTCTCCCGGGGTGGCCCAACTTGGGAAGACGGGGTT 881
Qy      299  erSerAspLeuIleuIle-----ArgLysCysSer 309
Db      882  CTTCCGGACCTCTCTTCTCTCTTTCGGGAATGCTCC 919

RESULT 5
CK000755
LOCUS   AGENCOURT_16363467 NIH_MGC_221 Homo sapiens cDNA clone
DEFINITION IMAGE:30707875 5', mRNA sequence.
ACCESSION CK000755
VERSION   CK000755.1 GI:38526789
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 797)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL
COMMENT   Contact: Daniela S. Gerhard, Ph.D.
          Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: James Martin, University of Iowa
          cDNA Library Preparation: M. Bento Soares, University of Iowa
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: NDAM1073 row: c column: 20
          High quality sequence stop: 656.
          Location/Qualifiers
FEATURES             source
     1..797
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:30707875"
     /lab_host="DH10B Tona"
     /clone_lib="NIH MGC 221"
     /note="Torgan: mixed; Vector: pYX-Asc; Site 1: EcoRI;
     Site 2: NotI; Library is oligo-dT primed and directionally
     cloned. Denatured RNA was size fractionated on a 1% agarose
     gel. First strand cDNA synthesis was primed with oligo-dT
     primer containing a Not I site. Double strand cDNA was
     size selected according to mRNA size fraction, ligated with
     EcoR I adaptor, digested with Not I and then cloned
     directionally into pYX-Asc vector. Average insert size
     4-5Kb. Adaptors 5' (AATCGGCACGAGG)3' and 5'd
     (CTCTGTCG)3'. 3' Linker sequence - GCGGCCGCTGAGAGCC T18.
     Sequencing primers 3' end: T3 promoter primer 5'd
     (ATTAAACCTCACTAAAGGA)3'. 5' End: T7 promoter primer 5'd
     (TATATACACTCACTATAGG)3'. Library was constructed in the
     laboratory of M. Bento Soares. Note: this is a NIH_MGC
     Library"
ORIGIN
Alignment Scores:
Pred. No.:      3 71e-131      Length:      797
Score:          1283.00      Matches:      249
Percent Similarity: 92.62%      Conservative: 2
Best Local Similarity: 91.88%      Mismatches: 8
Query Match:      74.72%      Indels:      12
DB:              7          Gaps:          2

US-10-631-958-2 (1-326) x CK000755 (1-797)

Qy      23  ArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleGlyAsnLys 42

```

```

Db      9   CGGCACGAGGCACCACTGTTTCCCTTAGCCTCCATCACCACCTGACATCATC----- 59
Qy      43  PheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThrLeu 62
Db      60  -----GTTTACTGAACATGCTAATCAGGCCAAGGAGACTCTG 95
Qy      63  TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 82
Db      96  TATGAGATTAACTAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGTATGTT 155
Qy      83  SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 102
Db      156  AGCGAGGTGCTGCCACGCTCTCATTTGGGAGGACGACAGGAGCGCGGGTCGACCAAAC 215
Qy      103  HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 122
Db      216  CACCCCGGGGTGCTGGTCCCGAGTAGCTCCGATTGGAATCATTCGCCGAGGGTCA 275
Qy      123  ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 142
Db      276  ACGGACTGCGTGTGTTACTCCACCGTGGGCACCGACGACGAGAAACCTCGGGCGTGCAT 335
Qy      143  ILeuValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 162
Db      336  ATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAACAGCACACTC 395
Qy      163  LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 182
Db      396  CTTGCGTACTCCGTTGCTTGGCTACGGCTTACGGGGACATCATCAAGGACAGT 455
Qy      183  GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 202
Db      456  GAGAAGAAACGGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTC 515
Qy      203  SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 222
Db      516  TCCCAACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAACAACACGCTGGATCT 575
Qy      223  ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 242
Db      576  CCAAGGATAGGAAGCCCTCGCGGACAGGATGCTTTGTCAGGCAAGCAAGCACAGCAG 635
Qy      243  LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 262
Db      636  CTGGAGGAGGAGCAGAGAAGACACTCTATGTTTGAAGCTCGCGGAGGACGTGGAGGAG 695
Qy      263  TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 282
Db      696  TGGNCAGTCTGCTGTGGGAAGTTTCTGGGCCNATCATGCACAAACATGTCTGTGCTTGT 755
Qy      283  ArgArgSerProArgGlyLeuSerProAlaAla 293
Db      756  CGCCGGAGCCCA---GGGGCTCTCCCGGCTGCC 785

RESULT 6
BQ057191
LOCUS   AGENCOURT_6769595 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5812382
DEFINITION AGENSEQ_5', mRNA sequence.
ACCESSION BQ057191
VERSION   BQ057191.1 GI:19816531
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 1047)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Lou Staudt

```


primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pVX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCACGAC. This library was created for the University Iowa Brain Anatomy Project (BNAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 2,44e-119 Length: 758
Score: 1176.00 Matches: 218
Percent Similarity: 95.55% Conservative: 18
Best Local Similarity: 88.26% Mismatches: 11
Query Match: 68.49% Indels: 0
DB: 6 Gaps: 0

US-10-631-958-2 (1-326) x CB246749 (1-758)

Qy 80 GlyMetPheSerGluValLeuHisGlyLeuLeGlyArgThrGlnArgSerAlaGlyVal 99
Db 1 GGCATGTTTCAGCGAGTCTGCATGGGTGATTTGGAGGAGCGACGAGCGCTGTATC 60
Qy 100 AspGlnAenHisProAtqAlaValLeuValProSerSerLeuArgGlyLeilePro 119
Db 61 GACCCCAATACCCCGGCGGTGCTGTGTCACAGTACCCCTCAGATCGGCATCATACC 120
Qy 120 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 139
Db 121 CGAGGTCACAGATTGTGTGTACTCAACAGTGGGCACAACAGCGCAGACATCG 180
Qy 140 AlaLeuHisLeileValGlyAspSerLeuAlaMetAspValSerSerValHisAen 159
Db 181 GCTTTGCACATCATTTATTTGGGACTCACTGCGCAATAGACGTCTCTGTGCACATACC 240
Qy 160 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIlele 179
Db 241 AACACGCTGCTCGGTACTCGGTCTCTCTGCTGGCTACGGTTCTCTACGGGACTTAATC 300
Qy 180 LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 199
Db 301 AAGGACAGTAAAGAAACGGTGGATGGCTCGTCCGCTATGATTTCTCAGGGTTGAAG 360
Qy 200 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 219
Db 361 ACCTTTCTCTCTCAGTACTATGAGGGACACTGTCTCTCTCCAGCAGACACACG 420
Qy 220 ValGlySerProAtqAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 239
Db 421 GTGGGATCTCCACGGGAAATAAACCCTGCCGGCTGGGTGCTCTGTGTGAGGAGAGC 480
Qy 240 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 259
Db 481 AAGCAACAGCTGGAAGAGAGAGAGAGAAAGCCCTGTATGCTCGGAAACGCCGAGAA 540
Qy 260 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaLeuAenAlaThrAsnMetSer 279
Db 541 ATGGAAGAGTGGCAAGTGACATGTGGGAAGTTCTCTGCGCATCAATGCCACCAATGTCC 600
Qy 280 CysAlaCysArgAtqSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 299
Db 601 TGTGCTTGTCTCGAGGCGCTCGGGGCTGTGTCATTTGCCATTTCTGGAGATGGGTCT 660
Qy 300 SerAspLeuLeileArgLysCysSerArgPheAenPheLeuArgPheLeuileArg 319
Db 661 TCTGACCTCATCTTATCCGGAAGTCTCCAGGTTCCAGTTCCTCCTGAGATTCTCATCCG 720
Qy 320 HisThrAenGlnAspGln 326
Db 721 CACACGAACCGAGGAGGACG 741

RESULT 8

BQ063738

LOCUS

DEFINITION BQ063738 1078 bp mRNA linear EST 02-APR-2002
AGENCOURT_6873251 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925382

ACCESSION

BQ063738

VERSION

BQ063738.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (Bases 1 to 1078)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2096 row: p column: 23
High quality sequence stop: 540.

FEATURES

Location/Qualifiers

1..1078

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5925382"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_99"

/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 9.35e-108 Length: 1078
Score: 1073.50 Matches: 219
Percent Similarity: 87.16% Conservative: 5
Best Local Similarity: 85.21% Mismatches: 15
Query Match: 62.52% Indels: 18
DB: 5 Gaps: 5

US-10-631-958-2 (1-326) x BQ063738 (1-1078)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIle 20

Db 121 CCAAAGCATTTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACAGGCAAGCGGATA 180

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40

Db 181 TATGAAGAAGAAAGTGGCCACCTGTTCCCTTAGCTCCATCACCACATGACATCATC--- 237

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60

Db 238 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 267

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysGly 80

Db 268 ACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTCTGTGTGCGGAGATGGT 327


```

Db      372 AAAGCACTGTATGTTTGGAAAGCTCGCGAGACGTCGGAGAGTGGCAAGTCGTCTGTGGG 431
Qy      269 LysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly 288
Db      432 AAGTTTCTGGCCATCAATGCCACAAACATGTCCTGTGTGTCGCGAGCCCCAGGGGC 491
Qy      289 LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCys 308
Db      492 CTCTCCCGGTCGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCCGGAATGC 551
Qy      309 SerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGln 326
Db      552 TCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCAG 605

RESULT 10
BP224560
LOCUS      BP224560 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA
DEFINITION clone DAT01067, mRNA sequence.
ACCESSION BP224560
VERSION    BP224560.1
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 581)
AUTHORS    Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE      Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
JOURNAL    Genome Res. 14 (9), 1711-1718 (2004)
COMMENT    Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES   source
            1..581
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DAT01067"
            /cell_type="lymphocyte"
            /cell_line="Daudi"
            /clone_lib="Sugano cDNA library, lymphocyte Daudi"
            /note="Burkitt's lymphoma"
ORIGIN
Alignment Scores:
Pred. No.:      2,4e-101      Length:      581
Score:          1012.00      Matches:    192
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.48%      Mismatches: 0
Query Match:    58.94%      Indels:    0
DB:             5           Gaps:      0

US-10-631-958-2 (1-326) x BP224560 (1-581)

Qy      66 AsnIleAspLysTyrAspGlyIleValCysValGlyAspGlyMetPheSerGluVal 85
Db      3 AACATAGACAATAACGACGGCATGCTGTGTGCGCGAGATGGTATGTTCCAGCGAGTG 62
Qy      86 LeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArg 105
Db      63 CTGCACGGTCTGATTGGGAGACGACAGAGAGCGCGGGGTGACACAGGACCCCGCG 122
Qy      106 AlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCys 125
Db      123 GCTGTGCTGTCCTCCCGATGAGCTCCGATTTGGATCATCTCCGCGAGGTCAACGGACTGC 182
Qy      126 ValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValVal 145

```

```

Db      183 GTGTGTTACTCCACCGTGGGCACGACGACGAGAAACCTCGCGCTGCATATCGTTGTT 242
Qy      146 GlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeuAlaTyr 165
Db      243 GGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACACAGCACACTCTCTTCGCTAC 302
Qy      166 SerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLys 185
Db      303 TCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGAAGAGAA 362
Qy      186 ArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHis 205
Db      363 CGGTGGTGTGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGACCTCTCTCTCCACCAC 422
Qy      206 CysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAsp 225
Db      423 TGCATATGAAGGGACAGTGTCTTCTCCCTCCCTGCACACACACGCTGGGATCTCCANGGAT 482
Qy      226 ArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGlu 245
Db      483 AGGAGCCCTGCCGGGCGAGATGCTTTGTTTGCAGGCAAGCAAGCAGCAGCTGGAGGAG 542
Qy      246 GluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu 258
Db      543 GAGCAGAAGAAAGCACTGTATGTTTGGAAAGCTGCGGAG 581

RESULT 11
BP310011
LOCUS      BP310011 Sugano cDNA library, brain Homo sapiens cDNA clone
DEFINITION NRR01274, mRNA sequence.
ACCESSION BP310011
VERSION    BP310011.1
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 584)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
            Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
JOURNAL    Genome Res. 14 (9), 1711-1718 (2004)
COMMENT    Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES   source
            1..584
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="NRR01274"
            /tissue_type="brain"
            /clone_lib="Sugano cDNA library, brain"
ORIGIN
Alignment Scores:
Pred. No.:      7,45e-98      Length:      584
Score:          980.50      Matches:    190
Percent Similarity: 93.66%      Conservative: 2
Best Local Similarity: 92.68%      Mismatches: 2
Query Match:    57.11%      Indels:    11
DB:             5           Gaps:      1

US-10-631-958-2 (1-326) x BP310011 (1-584)

Qy      10 ProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPhe 29
Db      1 CCGTTTGGAGAAAGACAGCGCGGATATATGAAAGAAAAGTGGCCACTCTGTC 60

```


Db 184 TACTCGTGTCCCTGCTGGCTACGCGCTTCTACGGGACATCATCAAGGACAGTGAAG 243

Qy 185 LysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHis 204

Db 244 AAACGGTGGTGGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCAC 303

Qy 205 HisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArg 224

Db 304 CACTGCTATGAAGGACAGTGTCTTCCCTCTGCACACACACGGTGGGATCTCCAAG 363

Qy 225 AspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGlu 244

Db 364 GATAGGAAGCCCTGCCGGGACAGATGCTTTGTTTCAGGCAAGCAAGCAGCTGGAG 423

Qy 245 GluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluTrpGln 264

Db 424 GAGGAGCAGAAGAAGCACTGTATGGTTTGGAAAGTGGCGAGAGCTGGAGGAGTGGCAA 483

Qy 265 ValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArg 283

Db 484 GTTCTGTGGGAAGTTTCTGCCATCAATGCCAACCAACATGCTCTGTCTGTTCGC 540

RESULT 15

BU214294
LOCUS 603756937F1 CSEQCHN04 Gallus gallus cDNA clone ChEST669110 5', mRNA
DEFINITION 725 bp linear EST 25-NOV-2002
sequence.
ACCESSION BU214294
VERSION BU214294.1 GI:25390255
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 725)
BOARDMAN, P.E., SANZ-EZQUERRO, J., OVERTON, I.M., BURT, D.W., BOSCH, E.,
FONG, W.T., TICKLE, C., BROWN, W.R.A., WILSON, S.A. and HUBBARD, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .725
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST669110"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_lib="CSEQCHN04"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

FEATURES

source

(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Alignment Scores:
Pred. No.: 1,45e-91 Length: 725
Score: 925.00 Matches: 168
Percent Similarity: 84.84% Conservative: 39
Best Local Similarity: 68.85% Mismatches: 33
Query Match: 53.87% Indels: 4
DB: 5 Gaps: 1

US-10-631-958-2 (1-326) x BU214294 (1-725)

Qy 79 AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly 98

Db 3 CACGCGCATGTTTCAGTGAAGTGCATGGTCTCATTTGGAAGAATGCAGAGGACTCTGGC 62

Qy 99 ValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIle 118

Db 63 ATAGACCAAAATAATCCCAAGCACCGTTAGTCAGTGCATATATAAGGATTTGGCATAAT 122

Qy 119 ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr 138

Db 123 CCTGCTGGCTCAACAGATGGGTATGCTATTCACACTGTTGGCATTTCTGATCCAGTAACA 182

Qy 139 SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHis 158

Db 183 TCAGCTCTTCATATTATTATAGTGACTGTGCAGCCTTTAGATGCTCATCTGTGCATCAA 242

Qy 159 AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 178

Db 243 AACCAACAGTATTGGAAGTATGCTGATCATTTGTTGGGCTATGGTTTTTATGGAGATGT 302

Qy 179 IleLysAspSerGluLysLysLysArgTrpLeuAlaArgTyrAspPheSerGlyLeu 198

Db 303 TTGAAGACACTGAAGAAGAACGGTGGATCGGTCGATGAGATACGACTATTTCAGGCTTC 362

Qy 199 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 218

Db 363 AAGACTTTTCTCTCATCTACTATGAGGAACAATTTCTTTTCAACACGACGAAACAC 422

Qy 219 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 238

Db 423 ACCTGGGATCTCCACGAGATAAAGATAGCTGCAGAACAGGATGTTATATTGCAAGAA 482

Qy 239 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu 258

Db 483 AGTGAGCACAACCTGGCGAACAACGCAAGAGTGT-----GGATTCAACATGAAGAA 536

Qy 259 AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet 278

Db 537 GATGAAGAAGATGGAAGTTATTACCGGGAATTTCTAGCCATCAATGTAGTAAATATG 596

Qy 279 SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 298

Db 597 TGCTGTGCTGTCCACGAAGTCCAAAGGCTCTTTCCCGGCGAGCTCATTTTAGCAGATGGT 656

Qy 299 SerSerAspLeuIleLeuIleArg-LysCysSerArgPheAsnPheLeuArg-PheLeuI 318

Db 657 TCAGTGCACCTGATTTAGTTCGGAAAATGCTCCAGATTTGATTTTTCGCGTTTATCTTG 716

Qy 318 leArgHis 320

Db 717 TCAGACAT 724

Search completed: September 6, 2005, 20:29:11
Job time : 3154.86 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:25:44 ; Search time 60.6246 Seconds
(without alignments)
2753.633 Million cell updates/sec

Title: US-10-631-958-2
Perfect score: 1717
Sequence: 1 PKHLVFNPGGKGGKRI.....KSRPNFLRLHTNQDQ 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1640.5	95.5	537	1 CEK1 HUMAN	Q8tct0 homo sapien
2	1503.5	87.6	409	2 Q6ZPK5	Q6zpk5 mus musculu
3	1503.5	87.6	531	1 CEK1 MOUSE	Q8k4q7 mus musculu
4	1295	75.4	339	2 Q6NX59	Q6nx59 homo sapien
5	1140	66.4	485	2 Q6GMF3	Q6gmf3 xenopus lae
6	1117.5	65.1	572	2 Q6GLV1	Q6glv1 xenopus lae
7	620.5	36.1	410	2 Q7PRA8	Q7pra8 anopheles g
8	613	35.7	487	2 Q95S15	Q95s15 drosophila
9	613	35.7	687	2 Q9VNA6	Q9vna6 drosophila
10	502.5	29.3	533	2 Q91JU45	Q91ju45 arabidopsis
11	498.5	29.0	608	2 Q6USK2	Q6usk2 arabidopsis
12	486.5	28.3	549	2 Q9TZ11	Q9tzi1 caenorhabdi
13	455.5	26.5	700	2 Q6H6H1	Q6h6h1 oryza sativ
14	436	25.4	532	2 Q6UZF6	Q6ufz6 homo sapien
15	389	22.7	485	2 Q817L1	Q817l1 arabidopsis
16	365.5	21.3	1240	2 Q65419	Q65419 arabidopsis
17	321.5	18.7	586	2 Q949C3	Q949c3 oryza sativ
18	311.5	18.1	732	2 Q9PHL3	Q9phl3 arabidopsis
19	311.5	18.1	763	2 Q9LRB0	Q9lrb0 arabidopsis
20	307.5	17.9	480	2 Q8H350	Q8h350 oryza sativ
21	301	17.5	685	2 Q7Q1P4	Q7qip4 anopheles g
22	300	17.5	446	2 Q84S01	Q84s01 oryza sativ
23	292	17.0	748	2 Q7XN57	Q7xn57 oryza sativ
24	287.5	16.7	757	2 Q94HY9	Q94hy9 oryza sativ
25	287.5	16.7	757	2 Q7KCS9	Q7kcs9 oryza sativ
26	282	16.4	641	2 Q7VYX8	Q7vyx8 drosophila
27	268	15.6	661	2 Q9VZW0	Q9vzw0 drosophila
28	261	15.2	280	2 Q6ZP59	Q6zps9 homo sapien
29	244.5	14.2	624	2 Q8GKF9	Q8gkf9 dictyosteli
30	243	14.2	654	1 SPH2 MOUSE	Q9jia7 mus musculu
31	236	13.7	616	2 Q6AYB2	Q6ayb2 rattus norv

ALIGNMENTS

RESULT 1	CEK1_HUMAN	STANDARD;	537 AA.
ID	CEK1_HUMAN	STANDARD;	PRT;
AC	Q8TCT0; Q9BYB3; Q9UGE5;		
DT	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Ceramide kinase (EC 2.7.1.138) (Acylsphingosine kinase) (hCERK) (Lipid kinase 4) (LK4).		
GN	Name=CERK; Synonyms=KIAA1646;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	SEQUENCE FROM N.A., AND CHARACTERIZATION.		
RC	TISSUE=Leukemia;		
RX	MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;		
RA	Sugiura M., Kono K., Liu H., Shimizugawa T., Minekura H., Spiegel S., Kohama T.;		
RT	"Ceramide kinase, a novel lipid kinase. Molecular cloning and functional characterization."		
RL	J. Biol. Chem. 277:23294-23300(2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Van Veldhoven P.P.;		
RT	"A search for lipid kinases."		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;		
RX	Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M., Clump M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K., Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J., Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G., Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R., Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G., Evans K.L., Fey J.M., Fleming K., French L., Garner A.A., Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C., Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S., Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A., Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M., Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T., McClay J., McLaren S., McMurray A., Milne S.A., Mortimore B.J., Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T., Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T., Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L., Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M., Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L., Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L., Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,		

RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorwan A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H., I.,
 RA Wang Q., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Pan S., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chisoso S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
 RA Schreier P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
 RA Korf I., Bedell J.A., Hillier L.W., Marais E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
 RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Edelman L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
 RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tiliahun Y., Wright H.,
 RA "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 RT [4]
 RN
 RP SEQUENCE OF 57-537 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=21156230; PubMed=11258795;
 RA Hitosawa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.;
 RT "Identification of novel transcribed sequences on human chromosome 22
 by expressed sequence tag mapping.";
 RL DNA Res. 8:1-9(2001).
 CC -!- FUNCTION: Catalyzes specifically the phosphorylation of ceramide
 to form ceramide 1-phosphate. Acts efficiently on natural and
 analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),
 to a lesser extent on C2-ceramide and C6-dihydroceramide, but not
 on other lipids, such as various sphingosines.
 CC -!- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.
 CC -!- COFACTOR: Calcium and magnesium.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
 CC -!- TISSUE SPECIFICITY: High level expression in heart, brain,
 skeletal muscle, kidney and liver; moderate in peripheral blood
 leukocytes and thymus; very low in spleen, small intestine,
 placenta and lung.
 CC -!- MISCELLANEOUS: Optimal pH is 6.0-7.5.
 CC -!- SIMILARITY: Contains 1 DAGKC domain.
 CC -!- CAUTION: Ref.3 sequence differs from that shown due to erroneous
 gene model prediction. An additional exon may exist between amino
 acid positions 168 and 169.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AB079066; BAC01154.1; -;
 DR EMBL; AJ457828; CAD29884.1; -;
 DR EMBL; AL096766; CAB62977.1; ALT SEQ.
 DR EMBL; AL118536; -; NOT ANNOTATED_CDS.
 DR EMBL; AB051433; BAB3316.1; -;
 DR Genew; HGNC:19256; CERK.
 DR GO; GO:0000289; C:integral to membrane of membrane fraction; IDA.
 DR GO; GO:0004685; F:calcium/calmodulin-dependent protein kinase. . .; IDA.
 DR GO; GO:0001729; F:ceramide kinase activity; IDA.
 DR GO; GO:0000287; F:magnesium ion binding; IDA.
 DR GO; GO:0006672; F:ceramide metabolism; TAS.
 DR InterPro; IPR001206; DAGKC.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR011036; PH related.
 DR Pfam; PF00781; DAGK_cat; 1.
 DR ProDom; PD005043; DAGKC; 1.
 DR SMART; SM00046; DAGKC; 1.
 DR SMART; SM00233; PH; 1.
 KW Calcium; Kinase; Magnesium; Transferase.

FT DOMAIN 132 278 DAGKC.
 SQ SEQUENCE 537 AA; 59977 MW; 3DBFC0ED8D679F7F CRC64;
 Query Match 95.5%; Score 1640.5; DB 1; Length 537;
 Best Local Similarity 96.3%; Pred. No. 8.7e-135; Indels 11; Gaps 1;
 Matches 314; Conservative 1; Mismatches 0;
 Qy 1 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTTDIIGNKFFVNVYVEVITEHANOAKE 60
 Db 130 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTTDIIGNKFFVNVYVEVITEHANOAKE 178
 Qy 61 TLVEINIDKDYGVVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSLSRIGIIPA 120
 Db 179 TLVEINIDKDYGVVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSLSRIGIIPA 238
 Qy 121 GSTDCVCYSTVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSYSLSLGCGYFGVDIIK 180
 Db 239 GSTDCVCYSTVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSYSLSLGCGYFGVDIIK 298
 Qy 181 DSEKKRWGLARYDFSGIKTFLSHHCYEGTVSFPLPAQHTVGSPPDRKPCRCAGCFVCQSK 240
 Db 299 DSEKKRWGLARYDFSGIKTFLSHHCYEGTVSFPLPAQHTVGSPPDRKPCRCAGCFVCQSK 358
 Qy 241 QOLEEEOKKALYLEAAEDVEEWMQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
 Db 359 QOLEEEOKKALYLEAAEDVEEWMQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418
 Qy 301 DLILIRKCSRFNLFRLIRHTNOODQ 326
 Db 419 DLILIRKCSRFNLFRLIRHTNOODQ 444
 RESULT 2
 Q6ZPK5 PRELIMINARY; PRT; 409 AA.
 ID Q6ZPK5 PRELIMINARY; PRT; 409 AA.
 AC Q6ZPK5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE MKTAA1646 protein (Fragment).
 GN Name=mKIAA1646;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic tail;
 RX PubMed=14621295;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
 RA Saga Y., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 cDNAs identified by screening of terminal sequences of 500 mouse KIAA-homologous
 cDNAs randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:167-180(2003).
 DR EMBL; AK129416; BAC98226.1; -;
 DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
 DR GO; GO:0007205; P:protein kinase C activation; IEA.
 DR InterPro; IPR001206; DAGKC.
 DR InterPro; IPR001849; PH.
 DR Pfam; PF00781; DAGK_cat; 1.
 DR ProDom; PD005043; DAGKC; 1.
 DR SMART; SM00046; DAGKC; 1.
 FT NON TER 1
 SQ SEQUENCE 409 AA; 45697 MW; 810E815B37E26808 CRC64;
 Query Match 87.6%; Score 1503.5; DB 2; Length 409;
 Best Local Similarity 86.8%; Pred. No. 6e-123;
 Matches 283; Conservative 19; Mismatches 13; Indels 11; Gaps 1;
 Qy 1 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTTDIIGNKFFVNVYVEVITEHANOAKE 60
 Db 8 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTTDIIGNKFFVNVYVEVITEHANOAKE 56

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGTORSAGVDQNHPRVLPVPSLRIGIIPA 120
 DB 57 TLYEINTDSYDGIIVCGDGMFSEVLHGVIGRTQOSAGIDPNHPRVLPVPSLRIGIIPA 116
 QY 121 GSTDCVCYSTGTCTSDAETSALHIVVGDLSLMDVSSVHNSLTLRYSVSLGFGFYGDILK 180
 DB 117 GSTDCVCYSTGTNDTAETSALHIIIGDSLAIIVSSVHNTLRLRYSVSLGFGFYGDILK 176
 QY 181 DSEKKRWGLARYDFSGLTFLSHHCYECTVGFSLPQAHTVGSPPDRKPCRCAGFCVCRQSK 240
 DB 177 DSEKKRWGLVRYDFSGLTFLSHQYEGTFLPQAHTVGSPPDRKPCRCAGFCVCRQSK 236
 QY 241 QOLEEEOKKALYGLAEAEVDEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
 DB 237 QOLEEEOKKALYGLAEAEVDEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 296
 QY 301 DLILIRKSRFNLFLIRHTNQDQ 326
 DB 297 DLILIRKSRFNLFLIRHTNQDQ 322

RESULT 3
 CEK1_MOUSE
 ID CEK1_MOUSE STANDARD; PRT; 531 AA.
 AC Q8K4Q7;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Ceramide kinase (EC 2.7.1.138) (Acylphosphatase) (mCERK).
 GN Name=Cerk;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;
 RA Sugita M., Kono K., Liu H., Shimizugawa T., Minekura H., Spiegel S.,
 RA Kohama T.;
 RA "Ceramide kinase, a novel lipid kinase. Molecular cloning and
 RT functional characterization.";
 RL J. Biol. Chem. 277:23294-23300 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Heart, and Thymus;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyszynski-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Iehii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki K., Shibata K., Shingawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 CC -!- FUNCTION: Catalyzes specifically the phosphorylation of ceramide
 CC to form ceramide 1-phosphate. Acts efficiently on natural and
 CC analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),
 CC to a lesser extent on C2-ceramide and C6-dihydroceramide, but not
 CC on other lipids, such as various sphingosines (By similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.
 CC -!- COFACTOR: Calcium and magnesium (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
 CC similarity).
 CC -!- TISSUE SPECIFICITY: High level expression in heart, brain and
 CC testis; low expression in spleen, liver and lung; not detected in
 CC skeletal muscle.
 CC -!- DEVELOPMENTAL STAGE: Highly expressed at 7 dpc and decreases
 CC rapidly thereafter.
 CC -!- SIMILARITY: Contains 1 DAGKc domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB079067; BAC01155.1; -;
 CC EMBL; AK042077; BAC31157.1; -;
 CC EMBL; AK052269; BAC34908.1; -;
 CC MGD; MGI:2386052; Cerk.
 CC GO; GO:0000299; C:integral to membrane of membrane fraction; ISS.
 CC GO; GO:0004685; F:calcium/calmodulin-dependent protein kinase. .; ISS.
 CC GO; GO:0001729; F:ceramide kinase activity; ISS.
 CC GO; GO:0000287; F:magnesium ion binding; ISS.
 CC GO; GO:0006672; P:ceramide metabolism; ISS.
 CC InterPro; IPR001206; DAGKc.
 CC InterPro; IPR011036; PH-related.
 CC Pfam; PF00781; DAGK_cat; 1.
 CC ProDom; PD005043; DAGKc; 1.
 CC Domain; Kinase; Magnesium; Transferase.
 CC DOMAIN 132 278
 CC FT CONFLICT 378 378 V -> M (in Ref. 2).
 CC FT CONFLICT 467 467 D -> Y (in Ref. 2; BAC34908).
 CC SEQUENCE 531 AA; 59811 MW; 31FEC534C348AA0A CRC64;
 Query Match 87.6%; Score 1503.5; DB 1; Length 531;
 Best Local Similarity 86.8%; Pred. No. 8.4e-123;
 Matches 283; Conservative 19; Mismatches 13; Indels 11; Gaps 1;
 QY 1 PKHLVFNPFPGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANQAKE 60
 DB 130 PKHLVFNPFPGKGQKRIYERKVAFLFTLASITTEII-----ITEHANQAKE 178
 QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGTORSAGVDQNHPRVLPVPSLRIGIIPA 120
 DB 179 TLYEINTDSYDGIIVCGDGMFSEVLHGVIGRTQOSAGIDPNHPRVLPVPSLRIGIIPA 238
 QY 121 GSTDCVCYSTGTSDAETSALHIVVGDLSLMDVSSVHNSLTLRYSVSLGFGFYGDILK 180
 DB 239 GSTDCVCYSTGTNDTAETSALHIIIGDSLAIIVSSVHNTLRLRYSVSLGFGFYGDILK 298
 QY 181 DSEKKRWGLARYDFSGLTFLSHHCYECTVGFSLPQAHTVGSPPDRKPCRCAGFCVCRQSK 240
 DB 299 DSEKKRWGLVRYDFSGLTFLSHQYEGTFLPQAHTVGSPPDRKPCRCAGFCVCRQSK 358
 QY 241 QOLEEEOKKALYGLAEAEVDEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
 DB 359 QOLEEEOKKALYGLAEAEVDEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418
 QY 301 DLILIRKSRFNLFLIRHTNQDQ 326
 DB 297 DLILIRKSRFNLFLIRHTNQDQ 322

```

Db 419 DLILIRKCSRFNLFRLIRHTNQEDQ 444
RESULT 4
Q6NX59
ID Q6NX59 PRELIMINARY; PRT; 339 AA.
AC Q6NX59
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE CERK protein.
GN Name=CERK;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC067255; AAH67255.1; -.
DR GO: GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO: GO:0007205; P:protein kinase C activation; IEA.
DR InterPro: IPR001206; DAGKC.
DR ProDom: PD005043; DAGKC; 1.
SQ SEQUENCE 339 AA; 37780 MW; A4C2ACDFF2E6F3D0 CRC64;

Query Match 75.4%; Score 1295; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 8.4e-105;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 MFSEVLHGLIGRTQSAGVDQNHPRVLPVPSLSRIGIIPAGSTDCVCYSTVGTSDAETSA 140
Db 1 MFSEVLHGLIGRTQSAGVDQNHPRVLPVPSLSRIGIIPAGSTDCVCYSTVGTSDAETSA 60

Qy 141 LHIVVGDSLAMDVSSVHNSITLLRSYVSLGGLGYGFIIDKSEKKRWGLGLARYDFSGLKT 200
Db 61 LHIVVGDSLAMDVSSVHNSITLLRSYVSLGGLGYGFIIDKSEKKRWGLGLARYDFSGLKT 120

Qy 201 FLSHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQOOLEBEQKALYGLEAARDV 260
Db 121 FLSHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQOOLEBEQKALYGLEAARDV 180

Qy 261 EEWQVVCCKFLAINATNNSCARRSPRGLSPAHLGDGSSDLILIRKCSRFNLFRLIRH 320
Db 181 EEWQVVCCKFLAINATNNSCARRSPRGLSPAHLGDGSSDLILIRKCSRFNLFRLIRH 240

Qy 321 TNQDDQ 326

```

```

Db 241 TNQDDQ 246
|||||
RESULT 5
Q6GMF3
ID Q6GMF3 PRELIMINARY; PRT; 485 AA.
AC Q6GMF3
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE MGC81777 protein.
GN Name=MGC81777;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC074110; AAH74110.1; -.
DR GO: GO:0004143; P:diacylglycerol kinase activity; IEA.
DR GO: GO:0007205; P:protein kinase C activation; IEA.
DR InterPro: IPR001206; DAGKC.
DR ProDom: PD005043; DAGKC; 1.
DR SMART: SM00046; DAGKC; 1.
SQ SEQUENCE 485 AA; 54874 MW; E7C30C87AE52214F CRC64;

Query Match 66.4%; Score 1140; DB 2; Length 485;
Best Local Similarity 62.4%; Pred. No. 4.9e-91;
Matches 204; Conservative 58; Mismatches 53; Indels 12; Gaps 2;

Qy 1 PKHLVFINPGKGGCKRIYERKVLFTLASITTDIIGNKFVYVVEVITEHANOAKE 60
Db 157 PKYLLVINPYGKGRQIYETKVLPLFSAAGICADVI-----VTEYANHARD 205

```

```
QY 61 TLYEINIDKYDGIIVCGDGMSEVLHGLIGTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 206 NUYEWLEKIDYGVVCGDGIIFSEVLHGLIGRVRQSDVDHNNPNAHLUSQCNIRIGIIPA 265
QY 121 GSTDCVCYSTGTSDAETSAHLIIVGDSLAMDVSVVHNSITLLRYSVSLGFGFYGDIIK 180
DB 266 GSTDCICVATVGINDPETSALHIIIGDCQPLDVCVSHNKRITFLKYSVSLGFGFYGDVLK 325
QY 181 DSEKRWGLARYDPSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSK 240
DB 326 GTEKRWGLPARYDVSCKTFLTHHCYEGSVSFQPAKWLVGSPRDQTPCTSGCYICROSS 385
QY 241 QOLEEEOKKALYGLE-AAEDVEEQVCGKFLAINATNMSCACRSPRGLSPAHLGDGS 299
DB 386 KOLEEQKKTQVSGSHRGKDDWTITIGRFMAINAVSMSCACRPTPKGLSPAHLADGS 445
QY 300 SDLIIRKCSRNFNRLIRHTNQDQ 326
DB 446 ADLIIVKCSRLDFLRLIRHTSNKDQ 472

RESULT 6
Q6GLV1
ID Q6GLV1 PRELIMINARY; PRT; 572 AA.
AC Q6GLV1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC84197 protein.
GN Name=MGC84197;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2];
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3];
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC074350; AAH74350.1; --
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat.1.
DR ProDom; PD005043; DAGKc; 1.
DR SMART; SM00046; DAGKc; 1.
SQ SEQUENCE 572 AA; 64899 MW; E3D22DB2C93B5B77 CRC64;

Query Match 65.1%; Score 1117.5; DB 2; Length 572;
Best Local Similarity 62.0%; Pred. No. 5.6e-89;
Matches 202; Conservative 54; Mismatches 59; Indels 11; Gaps 1;

QY 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDILGNKFYVYVEVITEHANOAKE 60
DB 157 PKNLVFINPFGGKGKQKIYENKVAFLFSAAGICADVI-----VTEYANHARD 205
QY 61 TLYEINIDKYDGIIVCGDGMSEVLHGLIGTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 206 HLYDANLEKIDYGVVCGDGMSEVLHGLIGRVRQSDVDHNNPNAHLUSQCNIRIGIIPA 265
QY 121 GSTDCVCYSTGTSDAETSAHLIIVGDSLAMDVSVVHNSITLLRYSVSLGFGFYGDIIK 180
DB 266 GSTDCICVATVGINDPETSALHIIIGDCQPLDVCVSHNKRITFLKYSVSLGFGFYGDVLK 325
QY 181 DSEKRWGLARYDPSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSK 240
DB 326 GSEKRWGLPARYDVSCKTFLTHHCYEGSVSFQPAKWLVGSPRDQTPCTSGCYICROSS 385
QY 241 QOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRSPRGLSPAHLGDGSS 300
DB 386 KOLEEQKKTQVCGSHRGKDDWTITIGRFMAINAVSMSCACRSPNGLSPAHLADGSA 445
QY 301 DLIIRKCSRNFNRLIRHTNQDQ 326
DB 446 DLIIVKCSRLDFLRLIRHTSNKDQ 471

RESULT 7
Q7PRA8
ID Q7PRA8 PRELIMINARY; PRT; 410 AA.
AC Q7PRA8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000010845 (Fragment).
GN Name=ENSANGP0000010845;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA801008859; EAA07556.2; --
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat.1.
DR ProDom; PD005043; DAGKc; 1.
FT NON TER 1
FT NON TER 410
SQ SEQUENCE 410 AA; 45669 MW; 79535E4C0D150154 CRC64;

Query Match 36.1%; Score 620.5; DB 2; Length 410;
Best Local Similarity 39.2%; Pred. No. 1.1e-45;
Matches 135; Conservative 53; Mismatches 119; Indels 37; Gaps 5;
```

```

Qy 1 PKHLLVFNPFGKGQGGKRIYERKVPALFTLASITTDIIGNKFYVNVYEVITEHANOAKE 60
Db 77 PKHLLLFNPGGKGNALALYERKVPALFRLAGVDINLI-----ITQRAQQIID 125
Qy 61 --TLYEINIDKYDGVCGGDMFSEVLHGLIGTORSAGVDQNHPRAVLVPSLRIGII 118
Db 126 IVTSKSLLDNDYGLVCCGGDTFAELFNGLVTRTMDMCGIDIKYP-AYLPKPNIPIGVI 184
Qy 119 PAGSTDCVCYVTGTSDAETSAHLHVCGDSLAMDVSSVHH-----N 159
Db 185 PAGSTDTVACCLNGTTDKTCIIHIIHQHGLDISAVYSDAAKCDEGASPACTGRPR 244
Qy 160 STLLRYSVSLGIFYGDIIDKSEKKRWLGARYDFGLKTLFSLHCHVCGTVSFLPAQHT 219
Db 245 POLLKLFASALSYGLVGLDIADYSEKRWMMGPKRYDYSFGFKFLANRGYNABIVHLDRRG 304
Qy 220 VGSFDRKPCRAGCFVCRSQKQLEBEOKKALYGLEAAEDVEEVOVCGKFLAINATMS 279
Db 305 KQDPNDGVRCCLKCARCKKAKYGRDCGGERASY-----EDDTEPLVVRGKFLMVSGANIS 360
Qy 280 CACRRSPRGLSPAHLGDSGLLIRKCSRNFNLFRLIRHTNQ 323
Db 361 CSCERSPOGFSYCHLGDGLLDLVLRHTSFNMLRLLLTWTISK 404

RESULT 8
Q95S15
ID Q95S15 PRELIMINARY; PRT; 487 AA.
AC Q95S15;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE HL01538p.
GN ORFNames=CG16708;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunco J., Paclab J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY061001; AAL28549.1; -.
DR FlyBase; FBgn0037315; CG16708.
DR GO; GO:0048102; P:autophagic cell death; IEP.
DR GO; GO:0035071; P:salivary gland cell death; IEP.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
DR SMART; SM00046; DAGKC; 1.
SQ SEQUENCE 487 AA; 53711 MW; A4E71EC40354BB07 CRC64;

Query Match 35.7%; Score 613; DB 2; Length 487;
Best Local Similarity 35.2%; Pred. No. 6.5e-45;
Matches 134; Conservative 59; Mismatches 108; Indels 80; Gaps 7;

Qy 2 KHLVFNPFPGKGQGGKRIYERKVPALFTLASITTDIIGNKFYVNVYEVITEHANOAKE 61
Db 4 RLLVFNPNPGGKAGATYERHVRPIFQLAGVDATCI-----TTQRAQVKDI 52
Qy 62 LYEINIDKYDGVCGGDMFSEVLHGLIGTORSAGVDQNHPRAVLVPSLRIGIIPAG 121
Db 53 LLSHDLGVYDAVCCVGGSGDTVAEVLNGLIFRQMRGLGLDEQRPPIPRP-ALPVGVIIPAG 111
Qy 122 STDCVCYVTGTSDAETSAHLHVCGDSLAMDVSSVHHNSTLLRYSVSLGFGFYDIIKD 181
Db 112 STDTIAYSMHGTAADVRTAAIHLVQHRGLDVCVSVSGQSLLRFFCASVLSYGLGDVAQ 171

```

RESULT 9

```

Q9VNA6
ID Q9VNA6 PRELIMINARY; PRT; 687 AA.
AC Q9VNA6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE CG16708-PA (CG16708-pb).
GN ORFNames=CG16708;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flocker A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K., Nixon K., Nuskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu J.C.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RN Science 287:2185-2195(2000).

```


DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Putative ceramide kinase.
 GN Name=P0519806.23; Synonyms=OJ1003_B06.39;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
 clone: P0519806.",
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005006; BAD25678.1; -;
 DR EMBL; AP004676; BAD25337.1; -;
 DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
 DR GO; GO:0007205; P:protein kinase C activation; IEA.
 DR InterPro; IPR001206; DAGKC.
 DR ProDom; PD005043; DAGKC; 1.
 KW Kinase.
 SQ SEQUENCE 700 AA; 77961 MW; FBB2C2A3DF31C75D CRC64;
 Query Match 26.5%; Score 455.5; DB 2; Length 700;
 Best Local Similarity 29.8%; Pred. No. 6.3e-31;
 Matches 134; Conservative 55; Mismatches 114; Indels 147; Gaps 16;
 QY 1 PKHLVFINPFGKGQKRIYERKVAFLTLASITTDIIGNKFFYNYVEVITEHANQAKE 60
 DB 182 PKSLMVFVHPGCGKGGCKNWE-TVAPLPERAKVTKVI-----VTQRAGHAYD 229
 QY 61 TLVEI---NIDKYDGIYCV----- 76
 DB 230 TLASLDKDLKFPDGVIAVNTINACLSLFDIKHNYKMSARENTLSYDPOSAAGHKSM 289
 QY 77 -----GDDGMFSEVL-----HGLIGRTORSAGVD--- 100
 DB 290 LIFYCFIINMKCOEHRNDLSNELTGDDANAISGSNTPDDEHPILSTTRSTGLDISS 349
 QY 101 -----QNHPRALV--PSS-LRIGIIPAGTDC-----VCYSTVGTSDAETSALHI 143
 DB 350 SDSDPCNGDQVLPVSPFNWFRLLIIPSGSTDAIVLSPDVC-STTGERDPVTSALLI 408
 QY 144 VVGDSLAMDVSV-----HHNSTLRYSVLLGYGYGDIIDKSEKRWLGLARYDF 195
 DB 409 ILGRISLDIAQVVRWKSPPSAEVSPTVRYAASFAGYGYGEVIRESEKYRMWGPARYDF 468
 QY 196 SGLKTFLSHHCYEGTVSFLPAOHT-----VGSPRDRKP-CRAGCFVC 236
 DB 469 SGTWVFLKRSYEAQVAFLENGVTHSLTASAEANNANGVQTLYQHNRIRKTCIRNCLIC 528
 QY 237 R---QSKQLEBEOKKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPA 293
 DB 529 KGTSTSEQNSEDENPDS---SRTACETPKWWSKGRFLSVGAIVISCRNAPDGLVADA 585
 QY 294 HLGSSDLILIRKCSRNFRLIRHTNQDQ 326
 DB 586 HLSDGFLHLLLRDCLP-----LPFYLWHLTQ 611
 RESULT 14
 Q6UZF6 PRELIMINARY; PRT; 532 AA.
 AC Q6UZF6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE Ceramide kinase-like protein.

GN Name=CERKL;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14681825;
 RA Tuson M., Marfany G., Gonzalez-Duarte R.;
 RT "Mutation of CERKL, a novel human ceramide kinase gene, causes
 RT autosomal recessive retinitis pigmentosa (RP26).";
 RL Am. J. Hum. Genet. 74:128-138(2004).
 DR EMBL; AY357073; AAR13670.1; -;
 DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
 DR GO; GO:0007205; P:protein kinase C activation; IEA.
 DR InterPro; IPR001206; DAGKC.
 DR Pfam; PF00781; DAGK_cat; 1.
 DR ProDom; PD005043; DAGKC; 1.
 KW Kinase.
 SQ SEQUENCE 532 AA; 59602 MW; C73E590F7C25EED1 CRC64;
 Query Match 25.4%; Score 436; DB 2; Length 532;
 Best Local Similarity 31.3%; Pred. No. 2.2e-29;
 Matches 103; Conservative 63; Mismatches 121; Indels 42; Gaps 6;
 QY 1 PKHLVFINPFGKGQKRIYERKVAFLTLASITTDIIGNKFFYNYVEVITEHANQAKE 60
 DB 166 PKSLKILLNPQSHKKEATQVYKVEPLLLKLAGIKTDV-----TIMEYEGHALS 214
 QY 61 TLVEINIDKYDGIYCVGGDMFSEVLHGLIGRTORSAGVDQNHPRALVLP--SSLRIGII 118
 DB 215 LKSCCEQGFDPGVVVCVGGDGSSEVAHALLLRAOKNAGMETDR---ILTPVRAQLPLGLI 271
 QY 119 PAGSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSVVHHNSTLIRYSVLLGYGYGDI 178
 DB 272 PAGSTNVLAHLHGVPVITATLHINGHVQLVDVCTFSTAGKLLRFGFSAM-FGFGRT 330
 QY 179 IKDSEKRWLG-LARYDPFSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVC 237
 DB 331 LALAEKYRWSFPNRRDFAVVVKALAKLKAEDCEISFLPFNSS----- 372
 QY 238 QSKQLEBEOKKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGD 297
 DB 373 -----DDVQRRAQGGPKSCDNDQMIOQFLNVSIMAIPLCLCSVAPRGLAPNRLNN 426
 QY 298 GSSDLILIRKCSRNFRLIRHTNQDQ 326
 DB 427 GSMALLIARTSPEFIFKHLKRYASVKNQ 455
 RESULT 15
 Q8L7L1 PRELIMINARY; PRT; 485 AA.
 ID Q8L7L1
 AC Q8L7L1;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein A4921540.
 GN Name=A4921540;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 15:12:23 ; Search time 7333.65 Seconds
(without alignments)
10664.096 Million cell updates/sec

Title: US-10-631-958-9

Perfect score: 1614

Sequence: 1 atgggggcacggggggcgc.....agccagactcacagctga 1614

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1614	100.0	1654	9	CR456404 Homo sapi
2	1614	100.0	2042	9	AJ457828 Homo sapi
3	1614	100.0	4413	6	AX457006 Sequence
4	1612.4	99.9	4432	6	AR541900 Sequence
5	1612.4	99.9	4445	9	AB079066 Homo sapi
6	1607.6	99.6	4463	6	BD102675 Ceramide
7	1447	89.7	4171	6	BD183468 Novel gen
8	1447	89.7	4171	9	AB051433 Homo sapi
9	1446.4	89.6	1840	6	AX224383 Sequence
10	1444	89.5	1772	9	BC067255 Homo sapi
11	1209.8	75.0	1459	6	CQ730476 Sequence
12	1120	69.4	2830	10	AB079067 Mus muscu
13	903	55.9	979	6	AX456998 Sequence
14	856.2	53.0	3661	10	AX129416 Mus muscu
15	656.8	40.7	1450	5	CR386590 Gallus ga
16	593.8	36.8	2494	5	BC074350 Xenopus 1
17	512.4	31.7	1520	5	BC074110 Xenopus 1
18	475	29.4	550	6	AX457005 Sequence
19	464.4	28.8	474	6	AX457001 Sequence

c	20	320.2	19.8	329	6	AX457002	Sequence
	21	312	19.3	522	6	AX224385	Sequence
	22	298	18.5	3039	9	BC004278	Homo sapi
	23	298	18.5	3074	9	HSM804105	Homo sapi
	24	243.6	15.1	564	6	AX718996	Sequence
c	25	211	13.1	70890	9	HSDA59H18	Human DNA
	26	211	13.1	181772	2	AC007689	Homo sapi
c	27	170.4	10.6	240483	2	AC095955	Rattus no
c	28	170.4	10.6	260331	2	AC135400	Rattus no
c	29	167.2	10.4	219077	2	AC116764	Mus muscu
	30	167	10.3	167	6	AX457003	Sequence
	31	153	9.5	153	6	AX457004	Sequence
	32	150	9.3	382	6	AX224387	Sequence
c	33	148.6	9.2	181772	2	AC007689	Homo sapi
c	34	144.2	8.9	89319	9	HSBK29F11	Human DNA
	35	126.6	7.8	1323	5	CR760136	Xenopus t
	36	124.6	7.7	2333	3	AK112750	Ciona int
	37	108.6	6.7	153756	5	BX248240	Zebrafish
c	38	106.2	6.6	325530	2	AC125699	Rattus no
c	39	100.6	6.2	1084	6	AR510195	Sequence
c	40	100.6	6.2	3178	3	BT014925	Drosophila
	41	99	6.1	1810	3	AY061001	Drosophila
	42	99	6.1	2064	6	AR509346	Sequence
	43	99	6.1	2173	6	CQ608040	Sequence
	44	99	6.1	10337	6	CQ608039	Sequence
c	45	99	6.1	22353	2	AC017269	Drosophila

ALIGNMENTS

RESULT 1
CR456404
LOCUS
DEFINITION Homo sapiens BK29F11.1 full length open reading frame (ORF) cDNA
ACCESSION CR456404.1 GI:47678338
VERSION CDNA; chromosome 22; ORF.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1654)
AUTHORS Collins, J.E., Wright, C.L., Edwards, C.A., Davis, M.P., Grinham, J.A., Cole, C.G., Goward, M.E., Aguado, B., Malliya, M., Moktab, Y., Huckle, E.J., Beare, D.M. and Dunham, I.
TITLE Direct Submision
JOURNAL Submitted (24-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: c22@sanger.ac.uk Manuscript
COMMENT Homo sapiens cDNA sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to isolate cDNA clones representing the full length open reading frame of well annotated protein coding genes on human chromosome 22. For more information see http://www.sanger.ac.uk/HGP/Chr22/.

FEATURES

source
1. 1654
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="22"
/clones="pGEM.bK29F11.1"
/lab_host="JM109"
1. 1654
/gene="bK29F11.1"
11. 1624
/gene="bK29F11.1"
/codon_start=1
/protein_id="CAG30290.1"
/db_xref="GI:47678339"
/translation="MGATGAAPLPDSVLWVKQRCVLSLEPARALLRWRSFGPGAGA
PGADACSPVPVSEIIAVETDVHGKQSGKQMKPYAFTVHCVKRRRRHWKQVQ

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Kossida, S. and Encinas, J.
TITLE	Regulation of human sphingosine kinase-like protein
JOURNAL	Patent: WO 0228906-A 9 11-APR-2002;
	Bayer Aktiengesellschaft (DE)
FEATURES	Location/Qualifiers
source	1..4413
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
CDS	/db_xref="taxon:9606"
	76..1689
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAD38256.1"
	/db_xref="GI:21715796"
	/translation="MGATGAEEPLQSLVWVKQORCAVLEPAPRALIRLWRRSPGPGAGA PGADACSPVEIIVAEVTEHQHGSGKQWKEPYATVHCVRARHRWKWAQV TFWPCPEQLCHLQTLREMLEKLTSPKHLVFINPFGKGQKRIYERKVAPLFTL ASITDILVTEHANQAKETLYEINIDKYDGVCGDGMFSEVLHGLIGRTQBSAGVD QNPRAVLVPSLRIGIPAGSDVCVSTVGTSDAETSAIHVFLSHGLSDAMDVSSVHH NSTLLRVSLGGLGVFGDITKDSKKRWLGLARYDFSLGTLFSLHSGTVEGTVFLPA QHTVSPDRKPRAGFCQKQSQLEEBQKALYCLEAAEDVESWVVCVKFLAIN ATMNSCADRSPRGLSPAAHLPGSSDLILIRKCSRNFNPLRIHTNQDQDFDTFV EYRVKKFOFTSKHMEDESDLKEGGKKRFGHICSSHPSCCTVSNSSNWCDEVLHS PAIEVRVHCOLVRLFARGIEENPKPSHS"
ORIGIN	
Query Match	100.0%; Score 1614; DB 6; Length 4413;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1614; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGGGGCGACGGGGCGGGGAGCCGCTGCAATCCGTGCTGTGGGTGAAGCAGCAGCGC 60
Db	76 ATGGGGCGACGGGGCGGGGAGCCGCTGCAATCCGTGCTGTGGGTGAAGCAGCAGCGC 135
Qy	61 TGGCCCTGAGCTGGAGCCGCGCGGCTCTGCTGGCTGTGGCGGAGCCCGGGGCC 120
Db	136 TGGCCCTGAGCTGGAGCCGCGCGGCTCTGCTGGCTGTGGCGGAGCCCGGGGCC 195
Qy	121 GGAGCCGGCCCCCGCGCGGATGCTGCTGTGTGCTGTATCTGAGATCATCGCGTT 180
Db	196 GGAGCCGGCCCCCGCGCGGATGCTGCTGTGTGCTGTATCTGAGATCATCGCGTT 255
Qy	181 GAGAAAACAGAGCTTACCGGAAACATCAAGCGCAGTGGAAAAATGGCAAAAATG 240
Db	256 GAGAAAACAGAGCTTACCGGAAACATCAAGCGCAGTGGAAAAATGGCAAAAATG 315
Qy	241 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGCAGCAGCCGCTGGAGTGGCG 300
Db	316 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGCAGCAGCCGCTGGAGTGGCG 375
Qy	301 CAGGTGACTTTCTGGTGTCCAGAGGACGCTGTGTCACTTGTGGCTGCAGACCTGCGG 360
Db	376 CAGGTGACTTTCTGGTGTCCAGAGGACGCTGTGTCACTTGTGGCTGCAGACCTGCGG 435
Qy	361 GAGATGCTGGAGAGCTGACGTCAGACCAAGCATTTTACTGGTATTATTCACACCGTTT 420
Db	436 GAGATGCTGGAGAGCTGACGTCAGACCAAGCATTTTACTGGTATTATTCACACCGTTT 495
Qy	421 GGAGGAAAAGCAAGCGACCGGATATATGAAAGAAAGTGGCACAACCTGTTACCTTA 480
Db	496 GGAGGAAAAGCAAGCGACCGGATATATGAAAGAAAGTGGCACCACCTGTTACCTTA 555
Qy	481 GCCTCCATCACCCTGACATCATCGTTACTGAACATGCTAATCAGGCCAAGGAGACTCTG 540
Db	556 GCCTCCATCACCCTGACATCATCGTTACTGAACATGCTAATCAGGCCAAGGAGACTCTG 615
Qy	541 TATGAGATTAAACATAGACAAATACGCGGCATCGTCTGTGTGGCGGAGATGGTATGTTTC 600

LOCUS AR541900 4432 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 148 from patent US 6743619.
ACCESSION AR541900
VERSION AR541900.1 GI:53933980
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 4432)
Tang,Y.T., Zhou,P., Goodrich,R., Liu,C., Asundi,V., Ren,F.,
Zhang,J., Zhao,Q.A., Yang,Y., Xue,A.J., Wehrman,T., Wang,J.-R.,
Wang,D. and Drmanac,R.T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6743619-A 148 01-JUN-2004;
FEATURES Location/Qualifiers
source 1..4432
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 99.9%; Score 1612.4; DB 6; Length 4432;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1613; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGGCGACGGGGCGGGCGGCGCTGCAATCGTCTGTGGTGAAGCAGCGC 60
Db 1204 ATGGGGGCGACGGGGCGGGCGGCGCTGCAATCGTCTGTGGTGAAGCAGCGC 1263

QY 61 TGGCGCGTGAGCTGGAGCGCGCGGCTCTGCTGCGCTGTGGCGAGCCCGGGCC 120
Db 1264 TGGCGCGTGAGCTGGAGCGCGCGGCTCTGCTGCGCTGTGGCGAGCCCGGGCC 1323

QY 121 GGAGCGCGCGCGCGCGCGGAGTGCCTGCTGTGCTGTATCTGAGATCATCGCGTT 180
Db 1324 GGAGCGCGCGCGCGCGGAGTGCCTGCTGTGCTGTATCTGAGATCATCGCGTT 1383

QY 181 GAGGAACAGAGCTTACAGCGGAACATCAAGCAGTGGAAATGGCAGAAATGGAAAG 240
Db 1384 GAGGAACAGAGCTTACAGCGGAACATCAAGCAGTGGAAATGGCAGAAATGGAAAG 1443

QY 241 CCTTAGCGCTTTACAGTTCACTGTGTAAAGAGACGACGCGCCGCTGGAAGTGGCG 300
Db 1444 CCTTAGCGCTTTACAGTTCACTGTGTAAAGAGACGACGCGCCGCTGGAAGTGGCG 1503

QY 301 CAGGTGACTTTCTGGTGTCCAGAGGAGCTGTGTCACTTGTGGTGCAGACCCCTGGG 360
Db 1504 CAGGTGACTTTCTGGTGTCCAGAGGAGCTGTGTCACTTGTGGTGCAGACCCCTGGG 1563

QY 361 GAGATGCTGGAGAGCTGACGTCAGACCAAGCATTTTACTGGTATTATCAACCGTTT 420
Db 1564 GAGATGCTGGAGAGCTGACGTCAGACCAAGCATTTTACTGGTATTATCAACCGTTT 1623

QY 421 GGAGGAAAAGGACAAGCGGCGGATATGAAAGAAAAGTGGCACCCTGTTCACCTTA 480
Db 1624 GGAGGAAAAGGACAAGCGGCGGATATGAAAGAAAAGTGGCACCCTGTTCACCTTA 1683

QY 481 GCCTCCATCACCCTGACATCATCGTTACTGAAACATGCTTAATCAGGCCAAGGAGCTCTG 540
Db 1684 GCCTCCATCACCCTGACATCATCGTTACTGAAACATGCTTAATCAGGCCAAGGAGCTCTG 1743

QY 541 TATGAGATTAAATAGACAAATACGACCGCATCGTCTGTGTCGGCGGAGATGGTATGTC 600
Db 1744 TATGAGATTAAATAGACAAATACGACCGCATCGTCTGTGTCGGCGGAGATGGTATGTC 1803

QY 601 AGCGAGTGTCTGCACCGTCTGATTGGGAGGAGCGAGAGAGCGCGGGTCCACCAAGAC 660
Db 1804 AGCGAGTGTCTGCACCGTCTGATTGGGAGGAGCGAGAGAGCGCGGGTCCACCAAGAC 1863

QY 661 CACCCCCGGGCTGTGCTGTGTCGCCAGTAGCCCTCCGGAATTGGAATCATTTCCCGCAGGGTCA 720
Db 1864 CACCCCCGGGCTGTGCTGTGTCGCCAGTAGCCCTCCGGAATTGGAATCATTTCCCGCAGGGTCA 1923

QY 721 ACGGACTGCGTGTGTACTCTCCACCGTGGGCGACCGAGCGCAGAAACCTCGCGCGTGCAT 780

Db 1924 ACGGACTGCGTGTGTGTACTTCCACCCTGGGACACGAGCGCAGAAACCTTCGGCGCTGCAT 1983

QY 781 ATCTGTTTGGGAGCTCGCTGGCCATGGATGTGTCTCAGTCCACCACACAGCACACTC 840

Db 1984 ATCTGTTTGGGAGCTCGCTGGCCATGGATGTGTCTCAGTCCACCACACAGCACACTC 2043

QY 841 CTTCCGCTACTCGTGTCTCGTCTGGGCTACCGGCTTCTACGGGGACATCATCAAGGACAGT 900

Db 2044 CTTCCGCTACTCGTGTCTCGTCTGGGCTACCGGCTTCTACGGGGACATCATCAAGGACAGT 2103

QY 901 GAGAAGAAACGGTGTGTGGGCTCTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCCTC 960

Db 2104 GAGAAGAAACGGTGTGTGGGCTCTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCCTC 2163

QY 961 TCCCAACCATGCTATGAAGGAGCAGTGTCTTCTCTCTGCAACACACACGCTGGGATCT 1020

Db 2164 TCCCAACCATGCTATGAAGGAGCAGTGTCTTCTCTCTGCAACACACACGCTGGGATCT 2223

QY 1021 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTTGGTGGAGCAAGCAAGCAGCAG 1080

Db 2224 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTTGGTGGAGCAAGCAAGCAGCAG 2283

QY 1081 CTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAGCTGCGGAGACGCTGGAGGAG 1140

Db 2284 CTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAGCTGCGGAGACGCTGGAGGAG 2343

QY 1141 TGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCGCACAAACATGTCTGTGCTTGT 1200

Db 2344 TGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCGCACAAACATGTCTGTGCTTGT 2403

QY 1201 CGCGGAGCCCGCAGGGGCTCTCTCCCGGCTGCGCCACTTTGGGAGACGGGTCTTTCTGACCTC 1260

Db 2404 CGCGGAGCCCGCAGGGGCTCTCTCCCGGCTGCGCCACTTTGGGAGACGGGTCTTTCTGACCTC 2463

QY 1261 ATCTCTATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAAC 1320

Db 2464 ATCTCTATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAAC 2523

QY 1321 CAGCAGGACCAAGTTTGAAGTTTCACTTTTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1380

Db 2524 CAGCAGGACCAAGTTTGAAGTTTCACTTTTGAAGTTTATCGCGTCAAGAAATTCAGTTT 2583

QY 1381 AGCTCGAAGCAGATGGAGGATAGAGACAGCAGCCTCAAGAGGGGGGGAAGAGCGCTTT 1440

Db 2584 AGCTCGAAGCAGATGGAGGATAGAGACAGCAGCCTCAAGAGGGGGGGAAGAGCGCTTT 2643

QY 1441 GGGCACATTTTCAGCAGCACCCCTCTCTGCTGTCACCGCTCTCCACAGCTTCTTGGAAC 1500

Db 2644 GGGCACATTTTCAGCAGCACCCCTCTCTGCTGTCACCGCTCTCCACAGCTTCTTGGAAC 2703

QY 1501 TGGCAGCGGAGGCTCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCATGCGCAGCTGGTT 1560

Db 2704 TGGCAGCGGAGGCTCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCATGCGCAGCTGGTT 2763

QY 1561 CGACTCTTTTGCAGGAGAAATTGAAGAGAAATCGAAGCCAGACTCACAGCTGA 1614

Db 2764 CGACTCTTTTGCAGGAGAAATTGAAGAGAAATCGAAGCCAGACTCACAGCTGA 2817

RESULT 5
AB079066
LOCUS Homo sapiens cerk mRNA for ceramide kinase, complete cds.
DEFINITION AB079066
ACCESSION AB079066
VERSION AB079066.1 GI:21624339
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sugiyama,M., Kono,K., Liu,H., Shimizuawa,T., Minekura,H.,

TITLE Spiegel, S. and Kohama, T.
Ceramide Kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND
FUNCTIONAL CHARACTERIZATION
J. Biol. Chem. 277 (26), 23294-23300 (2002)
PUBMED 11956206
REFERENCE 2 (bases 1 to 4445)
AUTHORS Sugira, M., Kono, K., Shimizugawa, T., Minekura, H., Spiegel, S. and
Kohama, T.
Direct Submission
Submitted (29-JAN-2002) Masako Sugira, Sankyo Co., Ltd.,
Pharmacology and Molecular Biology Research Laboratories; 2-58
Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan
(E-mail: maugira@shina.sankyo.co.jp, Tel: 81-3-3492-3131,
Fax: 81-3-5436-8565)

FEATURES

source

1..4445
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..4445
/gene="cerk"
124..1737
/gene="cerk"
/codon_start=1
/product="ceramide kinase"
/protein_id="BAC01154.1"
/db_xref="GI:21624340"
/translation="MGRTGAEPLOSVLVWVQKRCVSLRPARALLRWSPGCGA
PGDACSVPVSEIIATVETRELVHGHQSGKRWKPEYAFVTHVCKRARRHWRKQV
TFWCPBQLCHLQTLREMLEKLTSRPLKLVFINFGGKGQKRLIRKRVAPLFTL
ASITDIIIVTHANAKETLYEINIDKYDGVCGDGMFSEVLHGLIGRTQSAVD
QNHPRVLVSLRIGIIPAGSTDVCVSTVGTDAETSLHI VVGDSLAMDVSSVHH
NSTLLRVSLGLGYFGVDIIKDKSEKRWGLARVDFSLKTLPLSHHCYSGTVSFLPA
HTNVSRRDPRPCAGVCGQSKQLEEEKKALYGLEAEDVEEVQVVCGRFLAIN
AHTNSCARRRPGLSPAAHVGDSLLIRKCSRNFNLFRLIRHNQQDQDFTFV
EYVKKPQFTSKMEDESDLKEGKKRFGHICSHSPCCCTVSNSSNDCDGEVLHS
PAIEVRVHCOLVRLFARGIEENKPDHS"

gene

CDS

ORIGIN

Query Match 99.9%; Score 1612.4; DB 9; Length 4445;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1613; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 ATGGGGCGAGCGGGCGCGGCGGCGGCTGCAATCCGTGCTGGGTGAAGCAGCAGCGC 60
124 ATGGGGCGAGCGGGCGCGGCGGCGGCTGCAATCCGTGCTGGGTGAAGCAGCAGCGC 183
61 TGGCGCGTGAGCTGGAGCCCGCGGCGGCTGCTGCGCTGGTGGCGAGCCCGGGGCC 120
184 TGGCGCGTGAGCTGGAGCCCGCGGCGGCTGCTGCGCTGGTGGCGAGCCCGGGGCC 243
121 GGAGCGCGCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
244 GGAGCGCGCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303
181 GAGGAAACAGACGTTACGGGAAACATCAAGCAGTGGAAATGCGAAAAATGGAAAG 240
304 GAGGAAACAGACGTTACGGGAAACATCAAGCAGTGGAAATGCGAAAAATGGAAAG 363
241 CTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACCGCACCGCTGGAAGTGGCG 300
364 CTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACCGCACCGCTGGAAGTGGCG 423
301 CAGGTGACTTTTGGTGTCCAGAGGAGAGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 360
424 CAGGTGACTTTTGGTGTCCAGAGGAGAGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 483
361 GAGATGCTGGAGAGAGCTGACGCTCCAGACCAAGCAATTTACTGGTATTTTATCAACCCGTTT 420
484 GAGATGCTGGAGAGAGCTGACGCTCCAGACCAAGCAATTTACTGGTATTTTATCAACCCGTTT 543
421 GGAGAAAGGACCAAGGCAAGCGGATATATGAAAGAAAGTGGCAACCACTGTTTCACTTTA 480

Db 544 GGAGAAAGGACCAAGCGGATATATGAAAGAAAGTGGCACCACCTGTTTCACTTTA 603
Qy 481 GCCTCCATCACCCTGACATCATCTGTTTACTGAAATGAAATGAAATGAAATGAAATGAAATG 540
Db 604 GCCTCCATCACCCTGACATCATCTGTTTACTGAAATGAAATGAAATGAAATGAAATGAAATG 663
Qy 541 TATGAGATTAACTAGACAAATACGACGGCATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db 664 TATGAGATTAACTAGACAAATACGACGGCATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 723
Qy 601 AGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db 724 AGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 783
Qy 661 CACCCCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db 784 CACCCCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 843
Qy 721 ACAGGACTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db 844 ACAGGACTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 903
Qy 781 ATCGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 904 ATCGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 963
Qy 841 CTTGCGTACTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 964 CTTGCGTACTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1023
Qy 901 GAGAAGAAACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Db 1024 GAGAAGAAACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1083
Qy 961 TCCACCACTGTATGAGGAGACAGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Db 1084 TCCACCACTGTATGAGGAGACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1143
Qy 1021 CCAAGGATAGGAAGCCCTCGCGGCGAGGATGCTTTGTTGTCAGGCAAGCAAGCAGCAG 1080
Db 1144 CCAAGGATAGGAAGCCCTCGCGGCGAGGATGCTTTGTTGTCAGGCAAGCAAGCAGCAG 1203
Qy 1081 CTGAGGAGGAGCAGAAAGAACGACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db 1204 CTGAGGAGGAGCAGAAAGAACGACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1263
Qy 1141 TGGCAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 1264 TGGCAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1323
Qy 1201 CGCGGAGCCCGCGGGGCGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1260
Db 1324 CGCGGAGCCCGCGGGGCGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1383
Qy 1261 ATCTCTATCCGGAAATGCTCCAGGTTCATTTTCTGAGATTTTCTCATCAGGCACACCAAC 1320
Db 1384 ATCTCTATCCGGAAATGCTCCAGGTTCATTTTCTGAGATTTTCTCATCAGGCACACCAAC 1443
Qy 1321 CAGCAGACCAAGTTTGACTTCACTTTTGTGAAAGTTTATCGCGTCAAGAAATTTCCAGTTT 1380
Db 1444 CAGCAGACCAAGTTTGACTTCACTTTTGTGAAAGTTTATCGCGTCAAGAAATTTCCAGTTT 1503
Qy 1381 ACCTCGAAGCAGCATGGAGGATGAGCAGCAGCTCAAGGAGGGGGGGAAGAACGGCTTTT 1440
Db 1504 ACCTCGAAGCAGCATGGAGGATGAGCAGCAGCTCAAGGAGGGGGGGAAGAACGGCTTTT 1563
Qy 1441 GGGCACAATTTGACAGCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
Db 1564 GGGCACAATTTGACAGCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1623
Qy 1501 TGGCAGCGGAGGTCTCTGACAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560
Db 1624 TGGCAGCGGAGGTCTCTGACAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1683

```
QY 1561 CGACTCTTTTGCACGAGGAATTGAAGAGAAATCCGAAGCCAGACTCACAGCTGA 1614
Db 1684 CGACTCTTTTGCACGAGGAATTGAAGAGAAATCCGAAGCCAGACTCACAGCTGA 1737

RESULT 6
BD102675 4463 bp linear PAT 27-AUG-2002
LOCUS Ceramide kinase and DNA thereof.
DEFINITION BD102675
ACCESSION BD102675
VERSION BD102675.1 GI:22648249
KEYWORDS WO 0196575-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 4463)
JOURNAL Sugiyura, M., Kono, K. and Kohama, T.
Ceramide kinase and DNA thereof.
SANKYO CO LTD, MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA
COMMENT Patent: WO 0196575-A 1 20-DEC-2001;
OS Homo sapiens (human)
PN WO 0196575-A/1
PD 20-DEC-2001
PF 11-JUN-2001 WO 2001JP004889
PR 14-JUN-2000 JP ODP 178039
PI MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA
PC C12N15/54, C12N9/12, C12N1/21, C12Q1/48, C07K16/40, A61K31/7125, PC
A61K48/00,
PC A61P3/04, A61P3/10, A61P7/00, A61P25/00, A61P29/00, A61P31/18, PC
A61P35/00
CC Ceramide kinase and DNA thereof
FH Key Location/Qualifiers
FT CDS (124)..(1734).
FEATURES
source 1..4463
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.6%; Score 1607.6; DB 6; Length 4463;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1610; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGGGCGACGGGGCGGCGGAGCGCGTGCATCCGTGCTGTGGTGAAGCAGCAGCGC 60
Db 124 ATGGGGCGACGGGGCGGCGGAGCGCGTGCATCCGTGCTGTGGTGAAGCAGCAGCGC 183

QY 61 TCGCGCGTGAGCCTGGAGCCCGCGCGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 120
Db 184 TCGCGCGTGAGCCTGGAGCCCGCGCGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 243

QY 121 GGAGCGGGCCCCCGGGCGGATCGCTGCTGTGCTGTGCTGTGATCTGAGATCATCGCGTT 180
Db 244 GGAGCGGGCCCCCGGGCGGATCGCTGCTGTGCTGTGCTGTGATCTGAGATCATCGCGTT 303

QY 181 GAGGAAACAGAGCTTTCACGGGAAACATCAAGGACGTGGAAATGGCAAAATGGAAGA 240
Db 304 GAGGAAACAGAGCTTTCACGGGAAACATCAAGGACGTGGAAATGGCAAAATGGAAGA 363

QY 241 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 300
Db 364 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 423

QY 301 CAGGTGACTTTCTGTGTCCAGAGAGCAGCTGTGTCACTTGTGCTGCAGACCCCTGGG 360
Db 424 CAGGTGACTTTCTGTGTCCAGAGAGCAGCTGTGTCACTTGTGCTGCAGACCCCTGGG 483

QY 361 GAGATGCTGGAGAGCTCAGCTCCAGACCAAGCAATTTACTGGTATTTTCAACCCGTTT 420
Db 484 GAGATGCTGGAGAGCTCAGCTCCAGACCAAGCAATTTACTGGTATTTTCAACCCGTTT 543
```

```
QY 421 CGAGGAAAGGACAAGGCAAGCGGATATATGAAAGAAAAGTGCGACCACTGTTTCACTTA 480
Db 544 CGAGGAAAGGACAAGGCAAGCGGATATATGAAAGAAAAGTGCGACCACTGTTTCACTTA 603

QY 481 GCCTCCATCACCACCTGACATCATCGTTTACTGAAACATGCTAATCAGGCCCAAGGAGACTCG 540
Db 604 GCCTCCATCACCACCTGACATCATCGTTTACTGAAACATGCTAATCAGGCCCAAGGAGACTCG 663

QY 541 TATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGCTATGTTTC 600
Db 664 TATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGCTATGTTTC 723

QY 601 AGCGAGTGTGCAAGCTCTGATTTGGGAGGACGACGAGAGCGCCGGGGTGCACCAAGAAC 660
Db 724 AGCGAGTGTGCAAGCTCTGATTTGGGAGGACGACGAGAGCGCCGGGGTGCACCAAGAAC 783

QY 661 CACCCCGGGCTGTGCTGGTCCCAAGTAGCTCCCGGATCGGATTCGAATCATTTCCCGCAGGTC 720
Db 784 CACCCCGGGCTGTGCTGGTCCCAAGTAGCTCCCGGATTCGAATCATTTCCCGCAGGTC 843

QY 721 ACGGAATGCTGTGTTTACTCCACCGTGGGCAACCGACGACGAGAAACCTCGCGCTGCAT 780
Db 844 ACGGAATGCTGTGTTTACTCCACCGTGGGCAACCGACGACGAGAAACCTCGCGCTGCAT 903

QY 781 ATCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAGCACACTC 840
Db 904 ATCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAGCACACTC 963

QY 841 CTTCCGCTACTCCGCTGCTCGCTGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT 900
Db 964 CTTCCGCTACTCCGCTGCTCGCTGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT 1023

QY 901 GAGAGAAACCGTGTGGGTCTTGCACAGATACGACTTTTTCAGGTTTAAAGACCTTCTCCTC 960
Db 1024 GAGAGAAACCGTGTGGGTCTTGCACAGATACGACTTTTTCAGGTTTAAAGACCTTCTCCTC 1083

QY 961 TCCACCACTGCTATGAAGGAGACAGTGTCTTCTCCTCGTGCACCAACACACCGTGGATCT 1020
Db 1084 TCCACCACTGCTATGAAGGAGACAGTGTCTTCTCCTCGTGCACCAACACACCGTGGATCT 1143

QY 1021 CCAAGGATAGGAGCCCTCGCGGCGAGGATGCTTTTGTTCAGGCAAGCAAGCAGCAG 1080
Db 1144 CCAAGGATAGGAGCCCTCGCGGCGAGGATGCTTTTGTTCAGGCAAGCAAGCAGCAG 1203

QY 1081 CTGAGAGGAGCAGAGAAAGACATGTAATGGTGGAAAGCTGCGGAGACGTGGAGGAG 1140
Db 1204 CTGAGAGGAGCAGAGAAAGACATGTAATGGTGGAAAGCTGCGGAGACGTGGAGGAG 1263

QY 1141 TGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCGCAAAACATGCTCTGTGCTTGT 1200
Db 1264 TGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCGCAAAACATGCTCTGTGCTTGT 1323

QY 1201 CGCGGAGCCCGGAGGCGCTCCCGGCTGCGGCTGCGGAGACGGGTCTTTCAGCCTC 1260
Db 1324 CGCGGAGCCCGGAGGCGCTCCCGGCTGCGGCTGCGGAGACGGGTCTTTCAGCCTC 1383

QY 1261 ATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1320
Db 1384 ATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1443

QY 1321 CAGCAGACCAAGTTTGAATTCATTTTGTGAAGTTTATTCGCGTCAAGAAATTCAGTTT 1380
Db 1444 CAGCAGACCAAGTTTGAATTCATTTTGTGAAGTTTATTCGCGTCAAGAAATTCAGTTT 1503

QY 1381 ACGTGCAAGCAGCAGGAGATGAGACAGGACCTCAAGGAGGGGGGAGAGAGGCTTT 1440
Db 1504 ACGTGCAAGCAGCAGGAGATGAGACAGGACCTCAAGGAGGGGGGAGAGAGGCTTT 1563

QY 1441 GGGCAGATTTGACAGCAGCCCTCTCTGCTGTGACCGCTCTCAACAGCTCTCTGGAAC 1500
Db 1564 GGGCAGATTTGACAGCAGCCCTCTCTGCTGTGACCGCTCTCAACAGCTCTCTGGAAC 1623
```

Qy	1501	TGCACGGGAGGTCTTGCA CAGCCCTGCGCATCAGAGTTCAGATGCCACACTGGTT	1560
Db	1624	TGCATCGGGAGGTCTTGCA CAGCCCTGCCATCAGAGTTCAGATGCCACACTGGTT	1683
Qy	1561	CGACTCTTTGCACGAGGAATTGAAGAGAATCCGAAGCCAGACTCACACAGCTGA	1614
Db	1684	CGACTCTTTGCACGAGGAATTGAAGAGAATCCGAAGCCAGACTCACACAGCTGA	1737
RESULT 7			
BD183468			
LOCUS	BD183468	4171 bp DNA linear	PAT 17-JUN-2003
DEFINITION	Novel genes and proteins encoded by the genes.		
ACCESSION	BD183468		
VERSION	BD183468.1	GI:31875668	
KEYWORDS	JP 2002345492-A/181.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 4171)		
TITLE	Novel genes and proteins encoded by the genes		
JOURNAL	Patent: JP 2002345492-A 181 03-DEC-2002;		
COMMENT	KAZUSA DNA RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002345492-A/181 PD 03-DEC-2002 PF 26-FEB-2002 JP 2002049009 PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA PC C12N15/09, C07K14/47/A61K31/711, A61K38/00, A61K48/00, A61P25/00, A61P25/14, PC A61P25/18, A61P35/00, C12N15/00, A61K37/02 CC Novel genes and proteins encoded by the genes FH Key Location/Qualifiers FT CDS Location/Qualifiers (2)..(1444).		
FEATURES	source		
ORIGIN			
Query Match	89.7%;	Score 1447;	DB 6; Length 4171;
Best Local Similarity	100.0%;	Pred. No. 2,4e-313;	
Matches 1447;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	168	GATCATCGCCGTTGAGGAAACAGACGTTACGGGAAACATCAAGGCGAGTGGAAAATGGCA	227
Db	1	GATCATCGCCGTTGAGGAAACAGACGTTACGGGAAACATCAAGGCGAGTGGAAAATGGCA	60
Qy	228	GAAATGGNAAGCCTTACGCTTTTACAGTTTCATGTGTAAGAGAGCACCACGGCACCG	287
Db	61	GAAATGGAAAGCCTTACGCTTTTACAGTTTCATGTGTAAGAGAGCACCACGGCACCG	120
Qy	288	CTGGAAGTGGCGCAGGTGACTTTCTGTGTGCCAGAGGAGCAGCTGTCTCACTTGTGGCT	347
Db	121	CTGGAAGTGGCGCAGGTGACTTTCTGTGTGCCAGAGGAGCAGCTGTCTCACTTGTGGCT	180
Qy	348	GCAGACCTTCGGGAGATGCTGGGAGAAGCTGACGTCACGATCCAGACCAAAGCAATTTACTGGTATT	407
Db	181	GCAGACCTTCGGGAGATGCTGGGAGAAGCTGACGTCACGATCCAGACCAAAGCAATTTACTGGTATT	240
Qy	408	TATCAACCCGTTTGGAGAAAAGACAAGCGCAAGCGGATATATGAAAGAAAAGTGGCAC	467
Db	241	TATCAACCCGTTTGGAGAAAAGACAAGCGCAAGCGGATATATGAAAGAAAAGTGGCAC	300
Qy	468	ACTGTTACCTTAGCCTCCATCACCACATGACATCATCGTTACTGAAACATGCTAAATCAGGC	527
Db	301	ACTGTTACCTTAGCCTCCATCAACATGACATCATCGTTACTGAAACATGCTAAATCAGGC	360
Qy	528	CAAGGAGACTCTGTATGAGATTAAACATAGACAAATAACGACGGCATCGTCTGTGTCGGCG	587

RESULT 8
AB051433
LOCUS
DEFINITION
AB051433
ACCESSION
VERSION
AB051433.1
GI:13359166
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Hirosewa, M., Nagase, T., Murahashi, Y., Kikuno, R. and Ohara, O.
TITLE
Identification of novel transcribed sequences on human chromosome
22 by expressed sequence tag mapping
JOURNAL
DNA Res. 8 (1), 1-9 (2001)
PUBMED
21156230
REFERENCE
11258795
2 (bases 1 to 4171)
AUTHORS
Ohara, O., Nagase, T. and Kikuno, R.
TITLE
Direct Submission
JOURNAL
Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdmainfo@kazusa.or.jp,
URL: <http://www.kazusa.or.jp/huge>, Tel:81-438-52-3913,
Fax:81-438-52-3914)
Location/Qualifiers
1. .4171
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hk01650"
/note="vector:pBluescriptII SK plus"
1. .4171
/gene="KIAA1646"
1. .1447
/gene="KIAA1646"
/note="Start codon is not identified."
/codon_start=2
/product="KIAA1646 protein"
/protein_id="BAB33316.1"
/db_xref="GI:13359167"
/translation="IIAIVEETDVHKGSGKQWKEKPYAFTVHCVRARRHRKWA
QVTFWPEQLCHLQTLREMLEKLSRPKHLVFINPFGKGQKRIYERKVAPLF
VLASITDIIIVTEHANQAKETLYEINIDKYIVCGDGMFSEVLHGLIGRTORSAG
TDONHPRAVLVPSSLRIGIIPAGSDVCYSTVGTSDAETSALHIVGDSLAMDVSV
HNSITLLYSVLLGCGYGVGDIIDSEKRWLGLARYDFSGLTFLSHCHYEVTSLF
PAQHTVSPRDKPCRCAGFCVCRSQKLEBEQKALYGLEAAEDVEEWQVCGKFLA
INATNMSCACRRSPRLSPAAHLGDSSDLILIRKCSRNFRLFLIRHTNQDDPFT
FVEYVRVKFQTSKMEDESDLKEGKGKRGHICSSHPSCCCTVSNSSNWCDEVL
HSPAIEVRVHCOLVRLFARGIEENPKPDSHS"

FEATURES
source

gene

CDS

ORIGIN

Query Match 89.7%; Score 1447; DB 9; Length 4171;
Best Local Similarity 100.0%; Pred. No. 2.4e-313;
Matches 1447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 168 GATCATCGCGCTTGAGGAAACAGACGTTCCACGGGAAACATCAAGGCAGTGGAAATGGCA 227
DB 1 GATCATCGCGCTTGAGGAAACAGACGTTCCACGGGAAACATCAAGGCAGTGGAAATGGCA 60
QY 228 GAAATGGAAAGCCCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGACGACGCGCACCG 287
DB 61 GAAATGGAAAGCCCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGACGACGCGCACCG 120
QY 288 CTGGAAGTGGGCGCAGGTGACCTTCTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 347
DB 121 CTGGAAGTGGGCGCAGGTGACCTTCTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 348 CGAGACCCCTCGGGAGAGTCTGGAGAGCTGAGCTCCAGACCAAGCATTTACTGGTATT 407

DB 181 GCAGACCCCTCGGGAGATGCTGGAGAAAGCTGACGTCCAGACCAAGCATTTTACTGGTATT 240
QY 408 TATCAACCCGTTTGGAGGAAAAGCAAGGCAAGCGGATATATGAAAGAAAAGTGGCACC 467
DB 241 TATCAACCCGTTTGGAGGAAAAGCAAGGCAAGCGGATATATGAAAGAAAAGTGGCACC 300
QY 468 ACTGTTTCACTTACCTCCATCACCACCTGACATCATCTGTTACTGTAACATGCTTAATCAGC 527
DB 301 ACTGTTTCACTTACCTCCATCACCACCTGACATCATCTGTTACTGTAACATGCTTAATCAGC 360
QY 528 CAAGGAGACTCTGTATGAGATTAAACATAGACAAAATACGACGGCATCTGTTGTGGCGG 587
DB 361 CAAGGAGACTCTGTATGAGATTAAACATAGACAAAATACGACGGCATCTGTTGTGGCGG 420
QY 588 AGATGGTATGTTACGAGGCTGCTGACCGTCTGATTGGGAGGACGACGAGAGAGCGCCCG 647
DB 421 AGATGGTATGTTACGAGGCTGCTGACCGTCTGATTGGGAGGACGACGAGAGAGCGCCCG 480
QY 648 GGTGACACGAGACACCCCGGCTGTGCTGGTCCCGAGTAGCCTCCGGATTGGATCAT 707
DB 481 GGTGACACGAGACACCCCGGCTGTGCTGGTCCCGAGTAGCCTCCGGATTGGATCAT 540
QY 708 TCCCGCAGGCTCAACCGGACTGCTGTGTTACTTCCACCGTGGGACACGACGACGAGAAAC 767
DB 541 TCCCGCAGGCTCAACCGGACTGCTGTGTTACTTCCACCGTGGGACACGACGACGAGAAAC 600
QY 768 CTCGCGCTGCATATCGTGTGTTGGGACCTCGCTGGCCATGAGATGTTCTCAGTCCACCA 827
DB 601 CTCGCGCTGCATATCGTGTGTTGGGACCTCGCTGGCCATGAGATGTTCTCAGTCCACCA 660
QY 828 CAACGACACTCTCTCGCTACTCGTCTCCCTGCTGGGCTACGGCTTCTACGGGACAT 887
DB 661 CAACGACACTCTCTCGCTACTCGTCTCCCTGCTGGGCTACGGCTTCTACGGGACAT 720
QY 888 CATCAAGGACAGTGAAGAAACCGTGTGGTCTTCCAGATACGACTTTTCAAGTTT 947
DB 721 CATCAAGGACAGTGAAGAAACCGTGTGGTCTTCCAGATACGACTTTTCAAGTTT 780
QY 948 AAAGACTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCAACA 1007
DB 781 AAAGACTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCAACA 840
QY 1008 CACGCTGGATCTCCAAGGATAGAAACCTCGCGGCGAGGATGCTTTGTTGACGCA 1067
DB 841 CACGCTGGATCTCCAAGGATAGAAACCTCGCGGCGAGGATGCTTTGTTGACGCA 900
QY 1068 AAGCAAGCAGCAGCTGGAGGAGGACGAGAAAGACACTGTATGTTTGGAGCTGCGA 1127
DB 901 AAGCAAGCAGCAGCTGGAGGAGGACGAGAAAGACACTGTATGTTTGGAGCTGCGA 960
QY 1128 GGACGTGGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACAT 1187
DB 961 GGACGTGGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACAT 1020
QY 1188 GTCTGTGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCATCTGGGAGACGG 1247
DB 1021 GTCTGTGCTTGTGCGCGGAGCCCGAGGGGCTCTCTCCCGGCTGCCCATCTGGGAGACGG 1080
QY 1248 GTCTTCTGACCTCACTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCAT 1307
DB 1081 GTCTTCTGACCTCACTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCAT 1140
QY 1308 CAGGCACCAACACGACGAGGACCACTTCACTTTTGTGAAGTTTATCGCGTCAA 1367
DB 1141 CAGGCACCAACACGACGAGGACCACTTCACTTTTGTGAAGTTTATCGCGTCAA 1200
QY 1368 GAAATTCAGTTTACGTCGAAGCAATGGAGGATGAGGACGACGACCTCAAGGAGGGGG 1427
DB 1201 GAAATTCAGTTTACGTCGAAGCAATGGAGGATGAGGACGACGACCTCAAGGAGGGGG 1260
QY 1428 GAAGAAGCGCTTTGGGACATTTTGCAGCAGCCACCTCTCTGCTGCTCAGCTCTCCAA 1487
DB 1261 GAAGAAGCGCTTTGGGACATTTTGCAGCAGCCACCTCTCTGCTGCTCAGCTCTCCAA 1320

Qy	1488	CAGCTCCTGGAAC	TGCGACGGGAGGT	CCTTGCACAGCCCTGCCATCGAGGT	CAGAGTCCA	1547
Db	1321	CAGCTCCTGGAAC	TGCGACGGGAGGT	CCTTGCACAGCCCTGCCATCGAGGT	CAGAGTCCA	1380
Qy	1548	CTGCCAGCTGGT	TCGACTCTTTGCACGAGGAA	TTGAAGAGAA	TCCGAAGCCAGACTCACA	1607
Db	1381	CTGCCAGCTGGT	TCGACTCTTTGCACGAGGAA	TTGAAGAGAA	TCCGAAGCCAGACTCACA	1440
Qy	1608	CAGCTGA	1614			
Db	1441	CAGCTGA	1447			
RESULT 9						
LOCUS	AX224383	AX224383	1840 bp	DNA	linear	PAT 10-SEP-2001
DEFINITION	Sequence 5 from Patent WO0160990.					
ACCESSION	AX224383					
VERSION	AX224383.1 GI:15554633					
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Rastelli, L.					
TITLE	Novel sphingosine kinases					
JOURNAL	Patent: WO 0160990-A 5 23-AUG-2001;					
	Curagen Corporation (US) ; GENENTECH, INC. (US)					
FEATURES	Location/Qualifiers					
source	1..1840					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
ORIGIN						
Query Match 89.6%; Score 1446.4; DB 6; Length 1840;						
Best Local Similarity 97.8%; Pred. No. 4.8e-313;						
Matches 1490; Conservative 0; Mismatches 1; Indels 33; Gaps 1;						
Qy	124	GCCGCGCCCGCGCGGAT	CGCTCTGTGCTGTATCTGAGATCAT	CGCGTTGAG	183	
Db	1	GCGCGCGCCCGCGCGGAT	CGCTCTGTGCTGTATCTGAGATCAT	CGCGTTGAG	60	
Qy	184	GAACAGACGTT	CAAGGAAACATCAAGCAGTGGAAATGCAGAAATGGAAGGCT	243		
Db	61	GAACAGACGTT	CAAGGAAACATCAAGCAGTGGAAATGCAGAAATGGAAGGCT	120		
Qy	244	TACGCTTTTACAGTT	CACGTGTAAAGAGAGACGACGCGCACCGCTGGAAGTGGGCGCAG	303		
Db	121	TACGCTTTTACAGTT	CACGTGTAAAGAGAGACGACGCGCACCGCTGGAAGTGGGCGCAG	180		
Qy	304	GTGACTTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCCCTCGGGGAG	363			
Db	181	GTGACTTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCCCTCGGGGAG	240			
Qy	364	ATGCTGGAGAGCTGACGTCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTTGA	423			
Db	241	ATGCTGGAGAGCTGACGTCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTTGA	300			
Qy	424	GGAAAGGACCAAGCGGATATGAAGAAAAGTGGCACCACCTGTTTACCTTTAGCC	483			
Db	301	GGAAAGGACCAAGCGGATATGAAGAAAAGTGGCACCACCTGTTTACCTTTAGCC	360			
Qy	484	TCCATCACCACTGACAT	CATCG-----TTACT	510		
Db	361	TCCATCACCACTGACAT	CATCGTTAACTATGTAGAAAGTAAATTA	420		
Qy	511	GAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGC	570			
Db	421	GAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGC	480			

Qy	571	ATCGTCTGTGTGCGCGAGAT	TGTTT	CAGCGAGGTGCTGCACGCTCTGATTTGGGAGG	630	
Db	481	ATCGTCTGTGTGCGCGAGAT	TGTTT	CAGCGAGGTGCTGCACGCTCTGATTTGGGAGG	540	
Qy	631	ACGAGAGGAGCGCGGGGT	CGACCAAA	CCACCCCGGGCTGTGCTGCTCCCACTAGC	690	
Db	541	ACGAGAGGAGCGCGGGGT	CGACCAAA	CCACCCCGGGCTGTGCTGCTCCCACTAGC	600	
Qy	691	CTCCGGATTGGAAT	TCATT	CCCGAGGGTCAACGACTGCGTGTGTTACTCCACCGTGGGC	750	
Db	601	CTCCGGATTGGAAT	TCATT	CCCGAGGGTCAACGACTGCGTGTGTTACTCCACCGTGGGC	660	
Qy	751	ACCAGCGACGAGAAAC	CTCGCGCTGCATAT	CGTTGTTGGGACTCGCTGGCCATGGAT	810	
Db	661	ACCAGCGACGAGAAAC	CTCGCGCTGCATAT	CGTTGTTGGGACTCGCTGGCCATGGAT	720	
Qy	811	GTGTCTTCAGTCCACCA	CACACACACTCTTCTGCTACTCCGTGTCCCTGCTGGGCTAC	870		
Db	721	GTGTCTTCAGTCCACCA	CACACACACTCTTCTGCTACTCCGTGTCCCTGCTGGGCTAC	780		
Qy	871	GGCTTTCTACGGGGACAT	CATCAAGGACAGTGAGAAACCGTGTGGGTCTTTGCCAGA	930		
Db	781	GGCTTTCTACGGGGACAT	CATCAAGGACAGTGAGAAACCGTGTGGGTCTTTGCCAGA	840		
Qy	931	TACGACTTTTTCAGGTT	TAAAGACCTTCTCTCCCACTGCTATGAGAGGACAGTGTCC	990		
Db	841	TACGACTTTTTCAGGTT	TAAAGACCTTCTCTCCCACTGCTATGAGAGGACAGTGTCC	900		
Qy	991	TTCTCTCTGCAACAC	CACCGTGGGATCTCCAGGATAGGAGCCCTGCGGGCAGGA	1050		
Db	901	TTCTCTCTGCAACAC	CACCGTGGGATCTCCAGGATAGGAGCCCTGCGGGCAGGA	960		
Qy	1051	TGCTTTGTTGTCAGG	CAAAAGCAGCAGCTGGAGGAGCAGAGAAAGCACTGTAT	1110		
Db	961	TGCTTTGTTGTCAGG	CAAAAGCAGCAGCTGGAGGAGCAGAGAAAGCACTGTAT	1020		
Qy	1111	GGTTTGAAGCTGCGAGG	ACGTGGAGAGTGGCAAGTGTCTGTGGGAAATTTCTGGCC	1170		
Db	1021	GGTTTGAAGCTGCGAGG	ACGTGGAGAGTGGCAAGTGTCTGTGGGAAATTTCTGGCC	1080		
Qy	1171	ATCAATGCCACAAACAT	GTCTGTGTCGCGGAGCCCCAGGGGCTCTCCCCGGCT	1230		
Db	1081	ATCAATGCCACAAACAT	GTCTGTGTCGCGGAGCCCCAGGGGCTCTCCCCGGCT	1140		
Qy	1231	GCCCACTTGGGAGACGGG	CTTCTGACTCATCTCATCCGAAATGCTCCAGGTTCAAT	1290		
Db	1141	GCCCACTTGGGAGACGGG	CTTCTGACTCATCTCATCCGAAATGCTCCAGGTTCAAT	1200		
Qy	1291	TTTCTGAGATTTCTCAT	ATAGGCACACCAACAGCAGGACCAAGTTTGACTTTCATTTTGT	1350		
Db	1201	TTTCTGAGATTTTCTCAT	ATAGGCACACCAACAGCAGGACCAAGTTTGACTTTCATTTTGT	1260		
Qy	1351	GAAGTTTATCGCGTCA	AGAAATTCAGTTTACGTCGAGACACATGGAGGATGAGGACAGC	1410		
Db	1261	GAAGTTTATCGCGTCA	AGAAATTCAGTTTACGTCGAGACACATGGAGGATGAGGACAGC	1320		
Qy	1411	GACCTCAAGGAGGGGG	GAAGAGCGCTTTGGGCACATTTTCAGCAGCAGCCCTCTCTGC	1470		
Db	1321	GACCTCAAGGAGGGGG	GAAGAGCGCTTTGGGCACATTTTCAGCAGCAGCCCTCTCTGC	1380		
Qy	1471	TGCTGACCGGTCTCAA	CAGCTCTCGAACTGCGACGGGGAGGCTCTGCA	CAGCCCTGCC	1530	
Db	1381	TGCTGACCGGTCTCAA	CAGCTCTCGAACTGCGACGGGGAGGCTCTGCA	CAGCCCTGCC	1440	
Qy	1531	ATCGAGGTCAAGTCC	ACCTGCTTCGACTCTTTGACGAGGAATTTGAAGAGAT	1590		
Db	1441	ATCGAGGTCAAGTCC	ACCTGCTTCGACTCTTTGACGAGGAATTTGAAGAGAT	1500		
Qy	1591	CCGAAGCCAGACT	CACACAGCTGA	1614		
Db	1501	CCGAAGCCAGACT	CACACAGCTGA	1524		

RESULT 10	
BC067255	
LOCUS	
DEFINITION	BC067255 Homo sapiens ceramide kinase, mRNA (CDNA clone IMAGE:6185601), complete cds.
ACCESSION	BC067255
VERSION	BC067255.1 GI:45595582
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1772) Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Hopkins, R.F., Jordan, H., Moore, T., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Dietchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullay, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Hale, S., Garcia, A.M., Gay, L.J., Richards, S., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalek, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCE	2 (bases 1 to 1772)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (12-MAR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgabbs@mail.nih.gov Tissue Procurement: Dr. James R. Lupski cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Falkquist, Anca Petrescu, Anna Liisa Prabh, Parvaneh Saeedi, JR Santos, Angeliq Schnerch, Ursula Skalek, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAC Plate: 141 Row: d Column: 4 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21703365 This clone has the following problem: The cds is short compared to the longest cds in the locus. Location/Qualifiers 1..1772 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6185601" /tissue_type="Peripheral Nervous System, dorsal root
source	

gene	ganglion"
	/clone_lib="Lupski_dorsal_root_ganglion"
	/lab_host="DH10B"
	/note="Vector: pCMV-SPORT6"
	1..1772
	/gene="CERK"
	/note="synonyms: KIAA1646, hCERK, da59H18.3, DXFp434E0211, FLJ21430, FLJ23239, da59H18.2, LK4"
	/db_xref="LocusID:64781"
	579..1598
	/gene="CERK"
	/codon_start=1
	/product="CERK protein"
	/protein_id="AAH67255.1"
	/db_xref="GI:45595583"
	/db_xref="LocusID:64781"
	/translation="MFSEVLHGLIQRORSAGVQDNHPRAVLVPSSLRIGIIPAGSTD
	CVCYSTVGTSDAETSAHLIVVGDLSAMDVSVHNSHTLLRVSLGLGYFGVDIIKDS
	EKKWLGARVDESLATFLSHCHVECTVSLPAQHTVGSPRDRKPCRCAGFCVCRQSK
	QQLSEKQKALYGLAEADVEEVQVCGKFLAINATNMSCACRSPRGLSPAHLIGDG
	SSDILILKCSRFRFLRHLRHTNQDQDFTFVEVYRVKXFPQFTSKHMEDESDLKE
	GKKRFGHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPK
	PDSHS"
ORIGIN	
Query Match	89.5%; Score 1444; DB 9; Length 1772;
Best Local Similarity	95.4%; Pred. No. 1.4e-312;
Matches 1539; Conservative	0; Mismatches 0; Indels 75; Gaps 2;
Qy	1 ATGGGGGCGACGGGGCGGCGGCGCGCTGCAATCCGTGCTGGGTCAAGCAGCAGCGC 60
Db	60 ATGGGGGCGACGGGGGCGGCGGCGCGCTGCAATCCGTGCTGGGTCAAGCAGCAGCGC 119
Qy	61 TGCGCCGTGAGCCCTGAGCCCGCGCGGCTCTGCTGCGCTGCTGCGGAGCCCGGGGCC 120
Db	120 TGCGCCGTGAGCCCTGAGCCCGCGCGGCTCTGCTGCGCTGCTGCGGAGCCCGGGGCC 179
Qy	121 GGAGCCGGCGCCCGCGCGGAGTCCCTGCTCTGCTGCTGCTGATCTGAGATCATCCCGTT 180
Db	180 GGAGCCGGCGCCCGCGCGGAGTCCCTGCTCTGCTGCTGATCTGAGATCATCCCGTT 239
Qy	181 GAGGAAACAGACGTTTACGGGAAACATCAAGCGAGTGGAAATGCGACAAATGGAAGA 240
Db	240 GAGGAAACAGACGTTTACGGGAAACATCAAGCGAGTGGAAATGCGACAAATGGAAGA 299
Qy	241 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGACACACCGGCACCGCTGGAAGTGGCG 300
Db	300 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGACACACCGGCACCGCTGGAAGTGGCG 359
Qy	301 CAGTGACTTTTCTGGTGTCCAGAGAGCAGCTGTGTCACTTGTGGCTGCAGACCTCGCG 360
Db	360 CAGTGACTTTTCTGGTGTCCAGAGAGCAGCTGTGTCACTTGTGGCTGCAGACCTCGCG 419
Qy	361 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGTTATTTATCAACCCGTTT 420
Db	420 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGTTATTTATCAAC----- 473
Qy	421 GGAGAAAGACAAAGCAAGCGGATATATGAAAGAAAGTGGACCACTGTTTACACTTA 480
Db	474 -----AAAGGACAAAGCGAGCGGATATATGAAAGAAAGTGGACCACTGTTTACACTTA 528
Qy	481 GCCTCCATCACCCTGACATCATCGTTTCTGTAATCATGCTAATCAGGCCAAGGAGACTCTG 540
Db	529 GCCTCCATCACCCTGACATCATC----- 552
Qy	541 TATGAGATTACATAGACAAATACGACGGCATCTGTGTGCGGGAGATGGTATGTTTC 600
Db	553 -----GCATCTGTGTGCGGGAGATGGTATGTTTC 584
Qy	601 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGGAGCGCGCGGGTGCACAGAAC 660
Db	585 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGGAGCGCGCGGGTGCACAGAAC 644


```
Qy 661 CACCCCGGGCTGCTGGTCCCAGTAGCTCCCGAATTGGAATCATTTCCGCGAGGTGCA 720
Db 645 CACCCCGGGCTGCTGGTCCCAGTAGCTCCCGAATTGGAATCATTTCCGCGAGGTGCA 704

Qy 721 ACGACTGCGTGTGTTACTCACCGTGGGACAGCGACGAGAACTCGGCGCTGCAT 780
Db 705 ACGACTGCGTGTGTTACTCACCGTGGGACAGCGACGAGAACTCGGCGCTGCAT 764

Qy 781 ATCGTTGTGGGACTCGTGGCCATGGATGTGCTCCTCAGTCCACCAACAGCACACTC 840
Db 765 ATCGTTGTGGGACTCGTGGCCATGGATGTGCTCCTCAGTCCACCAACAGCACACTC 824

Qy 841 CTTCCGCTACTCCGTGCTCCGCTGCTGAGTCAAGGAGTCAAGGAGTCAAGGAGTCA 900
Db 825 CTTCCGCTACTCCGTGCTCCGCTGCTGAGTCAAGGAGTCAAGGAGTCAAGGAGTCA 884

Qy 901 GAGAGAAACGGTGGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC 960
Db 885 GAGAGAAACGGTGGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC 944

Qy 961 TCCACCACTCTATGAAGGACAGTGTCTTCTCCTGCAACACACCGGTGGATCT 1020
Db 945 TCCACCACTCTATGAAGGACAGTGTCTTCTCCTGCAACACACCGGTGGATCT 1004

Qy 1021 CCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGAG 1080
Db 1005 CCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGAG 1064

Qy 1081 CTGAGGAGGACAGAAAGCACTGTATGGTTTGGNAGCTGCGGAGGAGTGGAGGAG 1140
Db 1065 CTGAGGAGGAGCAAGAAAGCACTGTATGGTTTGGNAGCTGCGGAGGAGTGGAGGAG 1124

Qy 1141 TGGCAAGTCGCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTGCTGTGT 1200
Db 1125 TGGCAAGTCGCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTGCTGTGT 1184

Qy 1201 CGCCGAGCCCGAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTTCGACCTC 1260
Db 1185 CGCCGAGCCCGAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTTCGACCTC 1244

Qy 1261 ATCCTCATCCGGAATGCTCCAGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1320
Db 1245 ATCCTCATCCGGAATGCTCCAGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1304

Qy 1321 CAGCAGGACCACTGTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
Db 1305 CAGCAGGACCACTGTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1364

Qy 1381 ACGTCGAAGCACATGGAGGATGAGGACAGCGACCTCAAGAGGGGGGGAAGCGCTTTT 1440
Db 1365 ACGTCGAAGCACATGGAGGATGAGGACAGCGACCTCAAGAGGGGGGGAAGCGCTTTT 1424

Qy 1441 GGGCACATTTGAGGAGGACCCCTCTGCTGCTGTCACCGTCTCCACAGCTCTCGGAC 1500
Db 1425 GGGCACATTTGAGGAGGACCCCTCTGCTGCTGTCACCGTCTCCACAGCTCTCGGAC 1484

Qy 1501 TGGCAGGGGAGTCTCTCAGACGCTTCCATCGAGGTGAGGTCCACTGCGAGCTGGTT 1560
Db 1485 TGGCAGGGGAGTCTCTCAGACGCTTCCATCGAGGTGAGGTCCACTGCGAGCTGGTT 1544

Qy 1561 CGACTCTTTGACGAGGAATTAAGAGAAATCCGAGCCAGACTCACAGCTGA 1614
Db 1545 CGACTCTTTGACGAGGAATTAAGAGAAATCCGAGCCAGACTCACAGCTGA 1598
```

```
RESULT 11
CQ730476 LOCUS 1459 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 16410 from Patent WO02068579.
ACCESSION CQ730476
VERSION CQ730476.1 GI:42304409
KEYWORDS Homo sapiens (human)
SOURCE
```

```
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 16410 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1..1459
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 75.0%; Score 1209.8; DB 6; Length 1459;
Best Local Similarity 91.3%; Pred. No. 3.7e-260;
Matches 1347; Conservative 0; Mismatches 2; Indels 126; Gaps 1;

Qy 140 CGGATGCTGCTCTGTGCTGTATCTGAGATCATCGCGTTGAGGAAACAGACGTTCAAG 199
Db 111 CAGATGCTGCTCTGTGCTGTATCTGAGATCATCGCGTTGAGGAAACAGACGTTCAAG 170

Qy 200 GGAAACATCAAGGAGTGGAAATGGCAGAAATGGAAAGCCCTTACGCTTTTACAGTTC 259
Db 171 GGAAACATCAAGGAGTGGAAATGGCAGAAATGGAAAGCCCTTACGCTTTTACAGTTC 230

Qy 260 ACTGTGTAAGAGAGACGACGCGACCTGGAAGTGGCGCAGGTGACTTCTTGGTGTGTC 319
Db 231 ACTGTGTAAGAGAGACGACGCGACCTGGAAGTGGCGCAGGTGACTTCTTGGTGTGTC 290

Qy 320 CAGAGGAGCAGCTGTGTACCTTGTGGTGCAGACCTCGCGGAGATGCTGAGAGAGCTGA 379
Db 291 CAGAGGAGCAGCTGTGTACCTTGTGGTGCAGACCTCGCGGAGATGCTGAGAGAGCTGA 350

Qy 380 CGTCAGACCAAGACATTTACTGGTATTTATCAACCCCTTTGGAGGAAAGGACAAAGCA 439
Db 351 ----- 350

Qy 440 AGCGGATATGAAAGAAAGTGGCACCATGTTTCACTTTAGCCTCCATCACCAGTACA 499
Db 351 ----- 350

Qy 500 TCATCGTTTACTGAACTGCTAATCAGGCAAGGAGACTCTGTATGAGATTAAACATAGACA 559
Db 351 -----TTACTGAACTGCTAATCAGGCAAGGAGACTCTGTATGAGATTAAACATAGACA 404

Qy 560 AATAACGACGGCATCGTCTGTGTCGGCGAGATGTTTTCAGCGAGGTGCTGACCGTTC 619
Db 405 AATAACGACGGCATCGTCTGTGTCGGCGAGATGTTTTCAGCGAGGTGCTGACCGTTC 464

Qy 620 TGATTGGAGAGACGACGAGGAGCGCGGGTCCAGCAGAACCAACCCCGGCTGTGCTGG 679
Db 465 TGATTGGAGAGACGACGAGGAGCGCGGGTCCAGCAGAACCAACCCCGGCTGTGCTGG 524

Qy 680 TCCCGAGTAGCTCCGAGTTGGAATCATTTCCCGCAGAGGTCAACCGACTGCGTGTACT 739
Db 525 TCCCGAGTAGCTCCGAGTTGGAATCATTTCCCGCAGAGGTCAACCGACTGCGTGTACT 584

Qy 740 CCACCGTGGGACCAAGGAGCGCAGAAACCTCGGCGCTGCAATCTGTTTGGGAGCTCGC 799
Db 585 CCACCGTGGGACCAAGGAGCGCAGAAACCTCGGCGCTGCAATCTGTTTGGGAGCTCGC 644

Qy 800 TGGCCATGGATGTGCTCTCAGTCCACCAACAGCAGACTCTTCTGCTACTCCGTGCTCC 859
Db 645 TGGCCATGGATGTGCTCTCAGTCCACCAACAGCAGACTCTTCTGCTACTCCGTGCTCC 704

Qy 860 TGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGAAGAGAAACGGTGGTTGG 919
Db 705 TGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGAAGAGAAACGGTGGTTGG 764
```


Qy	920	GTCTTGCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAG	979
Db	765	GTCTTGCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAG	824
	980	GGACAGTGTCTCTCTCTCCCTGCAACAACACGCGTGGGATCTCCAAGGGATAGGAAGCCCT	1039
Db	825	GGACAGTGTCTCTCTCTCCCTGCAACAACACGCGTGGGATCTCCAAGGGATAGGAAGCCCT	884
Qy	1040	GCCGGGACAGNATGCTTTTGTGTCAGGCAAAAGCAAGCAGCAGCTGGAGGAGGACGAGAAGA	1099
Db	885	GCCGGGACAGNATGCTTTTGTGTCAGGCAAAAGCAAGCAGCAGCTGGAGGAGGACGAGAAGA	944
Qy	1100	AAGCACTGTATGTGTTTGGAACTCCGGAGGACGTGGAGGAGTGCAAGTGGTCTGTGCGGA	1159
Db	945	AAGCACTGTATGTGTTTGGAACTCCGGAGGACGTGGAGGAGTGCAAGTGGTCTGTGCGGA	1004
Qy	1160	AGTTTCTGGCCCATCAATGCCCAAAACATGCTCTGTGTTGCGCCGAGCCCAAGGGGCC	1219
Db	1005	AGTTTCTGGCCCATCAATGCCCAAAACATGCTCTGTGTTGCGCCGAGCCCAAGGGGCC	1064
Qy	1220	TCTCCCGGCTGCCACCTTGGGACACGGGTCTCTGACCTCATCCTCATCCGGGAAATGCT	1279
Db	1065	TCTCCCGGCTGCCACCTTGGGACACGGGTCTCTGACCTCATCCTCATCCGGGAAATGCT	1124
Qy	1280	CCAGGTTCAATTTTCTGAGATTTCTCATAGGCACACCAACCAGCAGCAGCCAGTTTGACT	1339
Db	1125	CCAGGTTCAATTTTCTGAGATTTCTCATAGGCACACCAACCAGCAGCAGCCAGTTTGACT	1184
Qy	1340	TCACCTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGATTTTACGTGCAAGCACATGAGG	1399
Db	1185	TCACCTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGATTTTACGTGCAAGCACATGAGG	1244
Qy	1400	ATGAGCACAGCAGCTCTCAAGGAGGGGGGAAGAGCGCTTTGGSCACATTTGACAGAGCC	1459
Db	1245	ATGAGCACAGCAGCTCTCAAGGAGGGGGGAAGAGCGCTTTGGSCACATTTGACAGAGCC	1304
Qy	1460	ACCCTCTCTGTCTGTCAACCGTCTCCAACAGCTCCTTGAACTGCGACGGGAGGTCTCTGC	1519
Db	1305	ACCCTCTCTGTCTGTCAACCGTCTCCAACAGCTCCTTGAACTGCGACGGGAGGTCTCTGC	1364
Qy	1520	ACAGCCCTGCCATCGAGGTCAAGTCCACATGCCAGCTGGTTTCGACTCTTTGCAACGAGAA	1579
Db	1365	ACAGCCCTGCCATCGAGGTCAAGTCCACATGCCAGCTGGTTTCGACTCTTTGCAACGAGAA	1424
Qy	1580	TTGAAGAGAAATCCGAAGCCAGACTCAACACAGCTGA	1614
Db	1425	TTGAAGAGAAATCCGAAGCCAGACTCAACACAGCTGA	1459

RESULT 12	
AB079067	LOCUS
AB079067	2830 bp mRNA linear ROD 27-JUN-2002
DEFINITION	Mus musculus cerk mRNA for ceramide kinases, complete cds.
ACCESSION	AB079067
VERSION	AB079067.1 GI:21624341
KEYWORDS	.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1
REFERENCE	Sugiura,M., Kono,K., Liu,H., Shimizugawa,T., Minekura,H., Spiegel,S. and Kohama,T. Ceramide Kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND FUNCTIONAL CHARACTERIZATION J. Biol. Chem. 277 (26), 23294-23300 (2002) 11956206
JOURNAL	2 (bases 1 to 2830)
PUBLISHED	Sugiura,M., Kono,K., Shimizugawa,T., Minekura,H., Spiegel,S. and Kohama,T.
AUTHORS	Direct Submission
TITLE	Submitted (29-JAN-2002) Masako Sugiura, Sankyo Co., Ltd., Pharmacology and Molecular Biology Research Laboratories; 2-58
JOURNAL	

Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan
(E-mail:msugiyura@china.sankyo.co.jp. Tel:81-3-3492-3131,
Fax:81-3-5436-8565)

FEATURES

Location/Qualifiers

1. .2830

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

1. .2830

/gene="cerk"

1. .1596

/gene="cerk"

/codon_start=1

/product="ceramide kinases"

/protein_id="BAC01155.1"

/db_xref="GI:21624342"

/translation="MGAMGAAGAPLHSLVLMVKRRRCVSLPEARALLRWRRSPSPGPGSA
PGADARSLVLSIIIAVEEDKCEKHAASSGRWHKMNPFATVHVRVVRHRRWKWARY
TWSADAEQLCHLWQTLRGLSLTSRPHLLVFNPGKGQGGKRIYEKTVAPLFTL
ASHTTIIITEHANAQAKETLYENTDSDYGVICVGGDGMFSEVLHGVIGRTQOSAGID
PHPRAVLVPSTRLGIIIPAGSTDCVCYSTVGTNDYETSALHIIIGDSLAIIDVSSYGI
HNTLLRYSVSLGIVGYGDLIIKDSKCRNMGVLVRYDFGLKTFLSHQYYEGTLUSFLPA
QHTVSPRNKPCACRAGFCVROSKQOLEEEKALYGLENAEEVEWQVTCGKGLFPAI
ATNMSCACRSPGGLSGFPFAHLDGDSLLIIRKCSRFNRLIRLIRHTNQBDQDFFTV
EYVRVKFHFTKSHVEDNDNSKEQEKQFKGIKDRPSCCTCSASRNSWNCDCGEVMHS
PATEVRVHCVLRLPARGITEERS"

source

gene

CDS

ORIGIN

Query Match	69.4%	Score 1120;	DB 10;	Length 2830;
Best Local Similarity	81.3%;	Pred. No. 4.8e-240;		
Matches 1311; Conservative	0;	Mismatches 300;	Indels 1;	Gaps 1;
Qy	1	ATGGGGCGAGACGGGGCGGGGCGGAGCGCTGCAATCGTGTGGGTGAAGCAGCAGCGC	60	
Db	1			
Qy	61	TGCGCCGTGAGCTTGAGAGCCCGCGGGCTCTGCTCGCTGTGGCGGAGCCCGGGGCC	120	
Db	61			
Qy	121	GGAGCGGGCCCCCGCGGGATGCTCTCTGTGCGCTGTATCTGAGATCATCGCGGTT	180	
Db	121			
Qy	181	GAGGAAACAGAGCTTTCACGGGAAACATCAAGCGAGTGGAAAATGGCAGAAAATGGA	240	
Db	181			
Qy	241	CGTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGCACCGCTGGAGTGGCG	300	
Db	241			
Qy	301	CAGTGACTTTCGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCCCTGCGG	360	
Db	301			
Qy	361	GAGATGCTGAGAAAGCTGACGTCAGACCAAGCATTTTACTGTGATTTATCAACCCGTTT	420	
Db	361			
Qy	421	GGAGGAAAGGACAGCGCAAGCGGATATGAAAGAAAAGTGGCACCATGTTTCACCTTA	480	
Db	421			
Qy	481	GCCTCCATCACCATGCATCATCGTTTACTGAACTAGCAATGCTTAATCAGGCCAAGGAGACTTG	540	
Db	481			
Qy	541	TATGAGATTAAATAGACAAATACGACCGCATCGTGTGTGGCGGAGATGGTAGTTTC	600	
Db	541			
Qy	601	AGCGAGTGCTGCACGGTCTGATTGGGAGGACGACAGGAGCGCCGGGGTCGACACAGAAC	660	
Db	601			

[illegible]

RESULT 13

AX456998

LOCUS

DEFINITION

ACCESSION

LOCUS	AX45698	979 bp	DNA	linear	PAT 06-JUL-2002
DEFINITION	Sequence 1 from Patent WO0228906.				

QY 1134 GGAGAGTGGCAAGTCGTCTGTGGAGTTCCTGGCCATCATGCCCACAAACATGTCCTG 1193
DB 781 GGAGAGTGGCAAGTCGTCTGTGGAGTTCCTGGCCATCATGCCCACAAACATGTCCTG 840
QY 1194 TCTTGTGTGCGCGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTC 1253
DB 841 TCTTGTGTGCGCGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTC 900
QY 1254 TGACCTCATCTCCATCCGGAATGTCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 1313
DB 901 TGACCTCATCTCCATCCGGAATGTCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 960
QY 1314 CACCAACACGAGGACCAG 1332
DB 961 CACCAACACGAGGACCAG 979

RESULT 14
AKI29416
LOCUS Mus musculus mRNA for mKIAA1646 protein. linear ROD 21-NOV-2003
DEFINITION AKI29416
ACCESSION AKI29416
VERSION AKI29416.1 GI:37360495
KEYWORDS FLI, CDNA.
SOURCE Mus_musculus (house mouse)
ORGANISM Mus_musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Koseki, H., Hiraoka, S.,
Saga, Y., Nagase, T., Ohara, O. and Koga, H.
Prediction of the coding sequences of mouse homologues of KIAA
gene: III. the complete nucleotide sequences of 500 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
DNA Res. 10 (4), 167-180 (2003)
22977043
14621295
2 (bases 1 to 3661)
Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.
Direct Submission
Submitted (23-JUL-2003) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: mousekazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)
The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing: Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5' - &
3'-end one pass sequencing.
Location/Qualifiers
1. .3661
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mp01246"
/tissue_type="embryonic tail"
/note="vector:modified pBC SK+"
1. .3661
/gene="mKIAA1646"
41. .1231
/gene="mKIAA1646"
/note="CDS is predicted by in silico analysis. Start codon
is not identified.
Start codon is not identified."
/codon_start=2
/evidence=not_experimental
/product="mKIAA1646 protein"
/protein_id="BAC98226.1"
/db_xref="GI:37360496"
translation="FCFAASRPKHLVFINPBGKGQKRIYEKTVAPLFTLASTTTE
LIIEHANQAKETLYEINTSDYGVGVGGDMFSEVLHVGIGRTQQSAGIDPHNPA
VLVPSTLRIGIIPAGSTDCVYSTVTGNDATSEALHIIIGDSLAIIDVSSVHYHNTLLR

FEATURES
source
gene
CDS
QY 380 CGTCAGACCAAGCATTTACTGTGTTATTAACAACCGTTTGGAGGAAAAGCAAGCA 439
DB 15 CTTCAAGACCGAAGCATTTGCTGGTATTCATCAACCTTTTCGAGGGAAGAGTCAAGGCA 74
QY 440 AGCGGATATATGAAGAAAAGTGGCACCACTGTTTACCTTAGCCCTTCCATCACCACATGACA 499
DB 75 AGCGCATCTATGAAGAAAAGTGGCGCCTCTGTTTACCTTGGCTTCCATCCTACGAGAGA 134
QY 500 TCATCGTTACTGAACATGCTAATCAGGCAAGGAGACTCTGTATGAGATTAAACATAGACA 559
DB 135 TCATCATTTACAGAGCATGCGCAACCAAGCAAGGAGACTTTTATACGAGATCAACACAGACA 194
QY 560 AATACGACGGCATCTGTCGTGCGCGAGATGTTATGTTCCAGCGAGGTGCTGCACGGTC 619
DB 195 GCTATGATGGCATCTGTCGTAGTGGGACGGCATGTTCCAGCGAGGTGCTGCATGGGG 254
QY 620 TGATTGGGAGGACGACAGGAGCGCCGGGTGCGACAGAACCCCGCGGCTGTGCTGG 679
DB 255 TGATTGGGAGGACGACAGGAGCGCTGTATCGACCCCAATCACCCCGCGCTGCTGG 314
QY 680 TCCCGTAGTCTCGGATGGAATCATTCGCCAGAGGTCAACGAGACTGCGTGTGTTACT 739
DB 315 TGCCCGTAGTCTCGGATGGAATCATTCGCCAGAGGTCAACGAGACTGCGTGTGTTACT 374
QY 740 CCACCGTGGGACGACGAGCGCAGAAACCTCGGCGCTGCATATGTTGTTGGGAGACTGCG 799
DB 375 CAACAGTGGGACAAACGACGAGAGCATCGGCTTTGACATCATATTTGGGAGACTGCG 434
QY 800 TGGCCATGGATGTCTCTCAGTCACCAACACAGACACACTCTTCTCGCTACTCCGTGTCCTC 859
DB 435 TGGCAATAGAGCTGTCTCTGTCACCTACCAACAGCTGCTCGGCTACTCGGTTCTC 494
QY 860 TGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAGAGAAAAGTGGTGG 919
DB 495 TGCTGGGCTACGGCTTCTACGGGACATTAATCAAGGACAGCGGAAAAGAAAGTGGTGG 554
QY 920 GTCTTGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAG 979
DB 555 GCCTCGTCCGGTATGATTTCTCAGGGTTGAAGACCTTCTCTCATCAGTACTATGAAG 614
QY 980 GGACAGTGTCTCTCTCCTGCACAAACACACAGCGTGGGATCTTCAAGGATAGGAAGCCCT 1039
DB 615 GGACACTGTCTCTCTCCAGCACAGCACACAGCGTGGGATCTCCAGCGGACAAATAACCT 674
QY 1040 GCCCGGACGATGCTTTGTTGGAGGCAAGCAAGCAGCAGCTGAGAGAGAGAGAGA 1099
DB 675 GCCCGGCTGGGTGCTTCTGTCAGGACAGCAAGCAACAGCTGCGAAGAGAGAGA 734
QY 1100 AAGCACTGTATGTTTGGAGCTCGGAGGACGTGGAGGAGTGGCAAGTGGCAAGTGGGA 1159
DB 735 AAGCCCTGTATGGCTGGAGAACCGCGAGGAAGTGGAAAGTGGCAAGTGGGA 794
QY 1160 AGTTTCTGGCCATCAATGCGCAACCAATGTCTGTGTTGTGCGCGAGGCCCGAGGGGCC 1219
DB 795 AGTTCTGGCCATCAATGCGCAACCAATGTCTGTGTTGTGCGAGGCCCTGGGGGCC 854
QY 1220 TCTCCCGGCTGCCCATCTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAATGCT 1279
DB 855 TGTCCCATTTGCCCATCTGGGAGATGGGTCTTCTGACCTCATCTTATCCGGAAGTCT 914
QY 1280 CCAGGTTCAATTTCTGAGATTTCTCATCGGACACCAACAGCAGGACCAAGTTGACT 1339

YSVLLGVGFYGDLLIKDSEKRWMLVYDFSLKTELHSHOYEGTSLFLPAQHTVGS
PRDKPGRAGFCVQRKQSEBEKALYGLLENABEVMQVTRKQKFLALINATNMSC
ACPRSGPLSFAHLGSDSLILIRKSRFNFLRFLIRHNDQDFVTVEYVRVK
KPHFTSKHVEDENDSKEOKQKFKGCKIDRPSCTCSASRSSWNCDEVMHSPAIEVR
VHCQLVRLFARGIEEBS"
ORIGIN
Query Match 53.0%; Score 956.2; DB 10; Length 3661;
Best Local Similarity 81.4%; Pred. No. 5.9e-181;
Matches 1004; Conservative 0; Mismatches 228; Indels 1; Gaps 1;
QY 380 CGTCAGACCAAGCATTTACTGTGTTATTAACAACCGTTTGGAGGAAAAGCAAGCA 439
DB 15 CTTCAAGACCGAAGCATTTGCTGGTATTCATCAACCTTTTCGAGGGAAGAGTCAAGGCA 74
QY 440 AGCGGATATATGAAGAAAAGTGGCACCACTGTTTACCTTAGCCCTTCCATCACCACATGACA 499
DB 75 AGCGCATCTATGAAGAAAAGTGGCGCCTCTGTTTACCTTGGCTTCCATCCTACGAGAGA 134
QY 500 TCATCGTTACTGAACATGCTAATCAGGCAAGGAGACTCTGTATGAGATTAAACATAGACA 559
DB 135 TCATCATTTACAGAGCATGCGCAACCAAGCAAGGAGACTTTTATACGAGATCAACACAGACA 194
QY 560 AATACGACGGCATCTGTCGTGCGCGAGATGTTATGTTCCAGCGAGGTGCTGCACGGTC 619
DB 195 GCTATGATGGCATCTGTCGTAGTGGGACGGCATGTTCCAGCGAGGTGCTGCATGGGG 254
QY 620 TGATTGGGAGGACGACAGGAGCGCCGGGTGCGACAGAACCCCGCGGCTGTGCTGG 679
DB 255 TGATTGGGAGGACGACAGGAGCGCTGTATCGACCCCAATCACCCCGCGCTGCTGG 314
QY 680 TCCCGTAGTCTCGGATGGAATCATTCGCCAGAGGTCAACGAGACTGCGTGTGTTACT 739
DB 315 TGCCCGTAGTCTCGGATGGAATCATTCGCCAGAGGTCAACGAGACTGCGTGTGTTACT 374
QY 740 CCACCGTGGGACGACGAGCGCAGAAACCTCGGCGCTGCATATGTTGTTGGGAGACTGCG 799
DB 375 CAACAGTGGGACAAACGACGAGAGCATCGGCTTTGACATCATATTTGGGAGACTGCG 434
QY 800 TGGCCATGGATGTCTCTCAGTCACCAACACAGACACACTCTTCTCGCTACTCCGTGTCCTC 859
DB 435 TGGCAATAGAGCTGTCTCTGTCACCTACCAACAGCTGCTCGGCTACTCGGTTCTC 494
QY 860 TGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAGAGAAAAGTGGTGG 919
DB 495 TGCTGGGCTACGGCTTCTACGGGACATTAATCAAGGACAGCGGAAAAGAAAGTGGTGG 554
QY 920 GTCTTGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAG 979
DB 555 GCCTCGTCCGGTATGATTTCTCAGGGTTGAAGACCTTCTCTCATCAGTACTATGAAG 614
QY 980 GGACAGTGTCTCTCTCCTGCACAAACACACAGCGTGGGATCTTCAAGGATAGGAAGCCCT 1039
DB 615 GGACACTGTCTCTCTCCAGCACAGCACACAGCGTGGGATCTCCAGCGGACAAATAACCT 674
QY 1040 GCCCGGACGATGCTTTGTTGGAGGCAAGCAAGCAGCAGCTGAGAGAGAGAGAGA 1099
DB 675 GCCCGGCTGGGTGCTTCTGTCAGGACAGCAAGCAACAGCTGCGAAGAGAGAGA 734
QY 1100 AAGCACTGTATGTTTGGAGCTCGGAGGACGTGGAGGAGTGGCAAGTGGGA 1159
DB 735 AAGCCCTGTATGGCTGGAGAACCGCGAGGAAGTGGAAAGTGGCAAGTGGGA 794
QY 1160 AGTTTCTGGCCATCAATGCGCAACCAATGTCTGTGTTGTGCGCGAGGCCCGAGGGGCC 1219
DB 795 AGTTCTGGCCATCAATGCGCAACCAATGTCTGTGTTGTGCGAGGCCCTGGGGGCC 854
QY 1220 TCTCCCGGCTGCCCATCTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAATGCT 1279
DB 855 TGTCCCATTTGCCCATCTGGGAGATGGGTCTTCTGACCTCATCTTATCCGGAAGTCT 914
QY 1280 CCAGGTTCAATTTCTGAGATTTCTCATCGGACACCAACAGCAGGACCAAGTTGACT 1339

Db 915 CCAGGTTCAACTCTCTGAGATTTCTCATCCGGCACACGAAACAGGAGGACCGATTGCACT 974
Qy 1340 TCACCTTTGTTGAAGTTTATCGCGTCAAGAAATTCAGTTTACGTCGAAGCACATGAGG 1399
Db 975 TCACCTTTGTTGAAGTTTATCGAGTCAAGAAATTCACCTTCACGTCGAAGCACGTGAAG 1034
Qy 1400 ATGAGGACAGCGACTCAAGAGAGGGGGAGAGAGCGCTTTGGGCACATTTGCGAGCAGCC 1459
Db 1035 ACAGAGCAATGACTCGAAGGAACAAGAGAGCAGAAGTTTGGGAAGATCTGCAAGGACA 1094
Qy 1460 ACCCTCTGCTGTCGACCGCTTCCAAAGAGCTCTCGAACTGCGAGCGGGAGGTCTGTC 1519
Db 1095 GACCCCTCTGCACTTGTCTGAGCTCCAGAGCTCTCGAACTCTCGAACTGCGAGCGGGAATCATGC 1154
Qy 1520 ACAGCCCTGCGATCGAGGTGAGGTCCAGCTGCGAGCTGGTTCGACTCTTTGCGACGAGGAA 1579
Db 1155 ACAGCCGCGCAATGAGGTGAGGTGCGACTGCCAGCTGGTGGCGCTCTTTGCTCGGGGAA 1214
Qy 1580 TTGAAG-AGAATCGAAGCCAGACTCACAGC 1611
Db 1215 TCGAGGAAGACTCATAGCAAGAACCCCAAGC 1247

RESULT 15
CR386590
LOCUS CR386590 1450 bp mRNA linear VRT 05-APR-2004
DEFINITION Gallus gallus finished cdna, clone CHEST291d19.
ACCESSION CR386590
VERSION CR386590.1 GI:46239349
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 1450)
Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
Crining,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
Tickle,C. and Wilson,S.A.
Direct Submission
Submitted (05-APR-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickest@ms.unist.ac.uk
BSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cdna
sequencing project.
This sequence is from the
BSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cdna collection,
from a library constructed by Elizabeth Bosch. cdna was prepared
from RNA extracted from heads, normalised, and poly A-trimmed.
EcoRI-NotI cut cdna was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
coli DH10B.

FEATURES
source
1..1450
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="white Leghorn, HiseX"
/db_xref="taxon:9031"
/clone="CHEST291d19"
/clone_lib="CSEQCHN23"
/dev_stage="stage 22"

ORIGIN

Query Match 40.7%; Score 656.8; DB 5; Length 1450;
Best Local Similarity 68.2%; Pred. NO. 2.6e-136;
Matches 946; Conservative 0; Mismatches 432; Indels 10; Gaps 2;
Qy 202 AAACATCAAGCGAGTGGAAATGGCAGAAATGGAAAGCGCTTACGCTTTTACAGTTTCAAC 261
Db 1 AAACAGCGTAGTATTGGCAATGGCAAAAGATGGCAAGCCCGCTTTTACAGTTTAT 60
Qy 262 TGTGTAAGAGAGCACGACCGGACCGCTGGGAAGTGGCGGCGAGGTGACTTTCTGGTGTCCA 321

Db 61 TACGTGAAAAAAGCCCGAAATCACCGCTGGCGGTGCAGAGATGTGACGTTTGGTGTGCT 120
Qy 322 GAGGAGCAGCTGTGTCACTTGTGGTGCAGACCCCTGGGAGAGATGCTGGAGAGAGCTGACG 381
Db 121 GATGAGATCTTGTGTAAACCAAGTGGATACAGGCACT----GAATTACTTTGAATTCGACG 176
Qy 382 TCAGAGACAAAGCATTTTACTTGGTATTTTATCAACCCGCTTTGAGAGAAAGCAAGCAAG 441
Db 177 TGTAGACCAAGCAGTTTGTGTATTAATCCATATGAGGAGAAACGCAAGGGAAG 236
Qy 442 CGGATATATGAAGAAAGTGGCAACCATGTTTCACTTACCTTAGCTCCATCACCATGACATC 501
Db 237 AGGATTTATGAACAGAAAGTTGCTCCACTCTTCACTTGGCTTCTATTCTTACTGATGTT 296
Qy 502 ATGTTTACTGAACATGTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAA 561
Db 297 GTCGTAACATGAATGCAACCATGCTAAGGACAAATTAATTTGAAGTTTAATTTAAATAA 356
Qy 562 TAGCAGCGCATCGTCTGTGTCGCGAGAGATGTTTTCAGCGAGGTGCTGCACGCTCTG 621
Db 357 TATGATGCTGTTGTTGTTGGTGGGACGGCATGTTTCACTGAAGTATGATGATGCTC 416
Qy 622 ATTGGGAGAGCGCAGAGAGCGCCGGGTTCGACAGAAACCAACCCCGGGCTGTGCTGGTC 681
Db 417 ATTGGAAAGATGCAGAGGAGACTCTGGCATAGACCAAAATAATCCCAAGACCCGTTAGTC 476
Qy 682 CCAGTAGCCTCCGGATTGGAATCATTTCCCGAGGGTCAACGAGCTCGGTGTGTTACTTCC 741
Db 477 CAGTGAATATAAGGATTTGGCATAATTCCTGCTGGCTCAACAGATTCCGTTATGCTATCA 536
Qy 742 ACCGTGGGACACGAGCGACGAGAAACCTCGCGCTGCATATCTGTTTGGGGACTCGCTG 801
Db 537 ACTGTTGGCATTTCTGATCCAGTAACATCAGCTCTTCATATTAATTAAGTGTGACTGTG 596
Qy 802 GCCATGATGTGCTCCTCAGTCCACCAACAAGCACAATCTCCTTCGCTACTCGTGTCCCTG 861
Db 597 CTTTAGATGCTCATCTGTGCATCAAAACAACACGTTTGTGAAGTATGCTGTATCATTTG 656
Qy 862 CTGGGCTACCGCTTCTACGGGGCATCATCAAGGACAGTGAAGAAACGTTGTTGGGT 921
Db 657 TTGGGCTATGGTTTTATGAGAGATGTTTGAAGACAGTGAAGAAACGTTGATGGGT 716
Qy 922 CTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCCTCCCAACCTGCTGTGAAGG 981
Db 717 CCGATGAGATACGACTATTTCAGGCTTCAAGACTTTTCTTCTCATCTACTATGAAGGA 776
Qy 982 ACAGTGTCTTCTCCTGCAACAACAACCGTGGGATCTCCAGGGATAGGAAGCCCTGC 1041
Db 777 ACAATTTCTTTTCAACGAGCAAAACAACATGCGGATCTCCACGAGATAAAGATAGCTGC 836
Qy 1042 CGGGCAGGATGCTTTGTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAAGAA 1101
Db 837 AGAACAGGATGTTATATTTGCAAGAAAGTGAAGCAACTGGCAGAACACGCAAGAA- 895
Qy 1102 GCACTGTATGTTTGGAAAGCTGCGGAGAGCTGGAGAGTGGCAAGTCTGCTGGGAAG 1161
Db 896 -----GTGTGGATTCAACATGAAGAGATGAAGAAGATGGAAGGTTATTAAAGGGAAA 950
Qy 1162 TTTCTGCGCATCAATGCGCAACAACATGTCTGTGCTTGTGTCGCGGAGGCCCGAGGGCTC 1221
Db 951 TTTCTAGCCATCAATGTAGTAAATATGTGTGTGCTGTGCTGTCCACGAAGTCCAAAAAGGCTT 1010
Qy 1222 TCCCGGGCTGCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCTCATCCGGAATGCTCC 1281
Db 1011 TCACCGGAGCTCAATTTAGCAGATGGTTTACGTGACCTGATCTTAGTTCGGAATGCTCC 1070
Qy 1282 AGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACACGAGGAGCCAGTTTCACTTC 1341
Db 1071 AGATTTGATTTCTGCGTTATCTTGTGAGCATACAAACAAGATGACCAAGTTTGACTTC 1130
Qy 1342 ACTTTTGTGAAGTTTATCGGTCAAGAAATTCAGTTTACGTTAGCTGAAGACATCGAGAT 1401

Db 1131 TCATTTGTAGATGTTTATCGGCTGAAGAGTTTTCATTTTACATCAAAGCTTTCAGAAGAC 1190
QY 1402 GAGGACAGCGACCTCAAGAGGGGGGGAAGAGCGCTTTGGGCACATTTGCAGCAGCCAC 1461
Db 1191 AATGAAAGCAGTGTACACAGACATAGGAAAGAAACATTTTGGCCAGTTCTGCAGAGATCAT 1250
QY 1462 CCTCCTGTGTGTGACCGTCTCCAAACAGCTCCTGGAACGTGCGCGGGAGGTCTCTGCAC 1521
Db 1251 CCAGCTGTGTGTGCAATATTGCTAATAGCACTTGGAAATTGTATGGAGAAACCCCTGGAT 1310
QY 1522 AGCCCTGCCATCGAGGTGAGAGTCCACTGCCAGCTGGTTCCGACTCTTTGCACGAGGAATT 1581
Db 1311 AGCTCAGCAATTGAAGTGAGGGTTTCACTGCCAGTTAATGAAACTGTTTGCAGGGGAATC 1370
QY 1582 GAAGAGAA 1589
Db 1371 GAGGAAAA 1378

Search completed: September 5, 2005, 23:41:46
Job time : 7343.65 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 5, 2005, 15:04:18 ; Search time 929.933 Seconds
(without alignments)
10274.357 Million cell updates/sec

Title: US-10-631-958-9
Perfect score: 1614
Sequence: 1 atggggggcagggggcggc.....agccagactcacagctga 1614

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001s:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1614	100.0	4413	6 ABL40828	ABL40828 Human sph
2	1614	100.0	4429	12 ADJ96598	Adj96598 Human lip
3	1612.4	99.9	4432	8 ABX70921	Abx70921 Novel hum
4	1612.4	99.9	4445	13 ADP55247	Adp55247 Human PRO
5	1610.8	99.8	1740	8 ADA05679	Ada05679 Human NOV
6	1607.6	99.6	4463	6 ABA96945	Abas96945 Human cer
7	1586.2	98.3	1740	12 ADN62844	Adn62844 Human NOV
8	1448.2	89.7	4231	3 AAA50510	Aaa50510 Human sph
9	1446.4	89.6	1840	5 AAD14426	Aad14426 Human sph
10	1251	77.5	3975	8 ACC70838	Acc70838 Human Sph
11	1017.6	63.0	4702	13 ADS10370	Ads10370 Human the
12	931	57.7	2241	5 AAS77728	Aas77728 DNA encod
13	903	55.9	979	6 ABL40822	ABL40822 Human sph
14	753.8	46.7	2186	5 AAS77730	Aas77730 DNA encod
15	752.4	46.6	1570	5 AAS77731	Aas77731 DNA encod
16	598.8	37.1	753	3 AAC76031	Aac76031 Human ORF
17	569	35.3	817	6 ABO99499	Abg99499 Human cod
18	569	35.3	817	13 ABL11669	Adsl11669 Human the
19	547	33.9	547	3 AAC76592	Aac76592 Human ORF
20	475	29.4	550	6 ABL40827	ABL40827 Human sph

21	464.4	28.8	474	6 ABL40823	ABL40823 Human sph
22	401.8	24.9	426	5 AAS77727	Aas77727 DNA encod
23	362.8	22.5	411	5 AAS77729	Aas77729 DNA encod
24	320.2	19.8	329	6 ABL40824	ABL40824 Human sph
25	312	19.3	522	5 AAD14427	Aad14427 Partial r
26	243.6	15.1	564	8 ABT23453	Abt23453 Immune-re
27	211	13.1	15181	4 AAK65588	Aak65588 Human imm
28	211	13.1	15185	4 AAK65589	Aak65589 Human imm
29	167	10.3	167	6 ABL40825	ABL40825 Human sph
30	153	9.5	153	6 ABL40826	ABL40826 Human sph
31	150	9.3	382	5 AAD14428	Aad14428 Partial m
32	102.6	6.4	839	6 ABO34636	Abg34636 Oligonucl
33	102.6	6.4	839	6 ABO34637	Abg34637 Oligonucl
34	99	6.1	2173	4 ABL25705	ABL25705 Drosophil
35	99	6.1	10337	4 ABL25704	ABL25704 Drosophil
36	83	5.1	1774	12 ADQ88890	Adq88890 Novel hum
37	71.4	4.4	839	6 ABO34634	Abg34634 Oligonucl
38	71.4	4.4	839	6 ABO34635	Abg34635 Oligonucl
39	69.8	4.3	498	6 ABK33019	Abk33019 DNA encod
40	67.8	4.2	1857	6 ABK90199	ABK90199 CDNA enco
41	67.8	4.2	2380	5 AAS14817	Aasi4817 Human cDN
42	67.8	4.2	2380	6 ABL59533	ABL59533 Human sph
43	67.8	4.2	2380	10 ADE85298	Ade85298 Farnesyl
44	67.8	4.2	2380	10 ADE38416	Ade38416 Human pro
45	67.8	4.2	2380	12 ADQ15181	Adql15181 Human Can

ALIGNMENTS

RESULT 1
ABL40828
ID ABL40828 standard; cDNA; 4413 BP.
XX
AC ABL40828;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human sphingosine kinase-like protein encoding cDNA.
XX
KW Human sphingosine kinase-like protein; intracellular signalling; gene;
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.
XX
OS Homo sapiens.
XX
FH Key
CDS Location/Qualifiers
FT 1..1689
FT /*tag= a
FT /product= "sphingosine kinase-like protein"
FT /note= "see ABB07857"
FT 76..1689
FT /*tag= b
FT /product= "sphingosine kinase-like protein"
FT /note= "see ABB07856"
XX
WO200228906-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-EP011516.
XX
PR 06-OCT-2000; 2000US-0238005P.
XX
PR 23-AUG-2001; 2001US-0314113P.
XX
XX (FARB) BAYER AG.
XX
PI Kossida S, Encinas J;
XX
DR WPI; 2002-340094/37.
XX
DR P-PSDB; ABB07856, ABB07857.
XX
XX New reagent for modulating the activity of sphingosine kinase-like
PT

PT protein polypeptide or polynucleotide and treating cancer, asthma,
PT allergy, an autoimmune disease, or a central or peripheral nervous system
PT disorder.
XX

PS Claim 1; Fig 9; 120pp; English.

XX The invention relates to a human sphingosine kinase-like protein. The
CC polypeptide can be expressed by standard recombinant methodology. The
CC sphingosine kinase-like protein and gene can be used to regulate
CC intracellular signalling and consequently cell proliferation and
CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
CC peripheral nervous system disorders (e.g. Parkinson's disease). The
CC present sequence represents the human sphingosine kinase-like protein
CC encoding cDNA

XX Sequence 4413 BP; 1020 A; 1107 C; 1207 G; 1079 T; 0 U; 0 Other;

Query Match 100.0%; Score 1614; DB 6; Length 4413;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGGGCGACGGGGCGGGCGGAGCGCTGCAATCCGTGCTGGGTGAAGCAGCAGGC	60
Db	76	ATGGGGCGACGGGGCGGGCGGAGCGCTGCAATCCGTGCTGGGTGAAGCAGCAGGC	135
Qy	61	TGCGCGGTGAGCCTGGAGCCGCGGGCTCTGCTGGCTGGTGGGAGCCGGGGCC	120
Db	136	TGCGCGGTGAGCCTGGAGCCGCGGGCTCTGCTGGCTGGTGGGAGCCGGGGCC	195
Qy	121	GGAGCGGCGCCCGGGCGGGATGCCCTCTGTGCTGTATCTGAGATCATCGCGTT	180
Db	196	GGAGCGGCGCCCGGGCGGGATGCCCTCTGTGCTGTATCTGAGATCATCGCGTT	255
Qy	181	GAGGAAACAGACGTTACGGGAAACATCAAGCGAGTGGAAATGCGAGAAATGGAAG	240
Db	256	GAGGAAACAGACGTTACGGGAAACATCAAGCGAGTGGAAATGCGAGAAATGGAAG	315
Qy	241	CCTTACGCTTTTACAGTTCTACTGTAAAGAGACGACGCGGACCGCTGGAAGTGGCG	300
Db	316	CCTTACGCTTTTACAGTTCTACTGTAAAGAGAGCAGCGGACCGCTGGAAGTGGCG	375
Qy	301	CAGGTGACTTTCTGCTGCCAGAGGAGCAGCTGTCTCTTGTGGCTGACAGCCCTGCG	360
Db	376	CAGGTGACTTTCTGCTGCCAGAGGAGCAGCTGTCTCTTGTGGCTGACAGCCCTGCG	435
Qy	361	GAGATGCTGGAGAGCTGACGCTCAGACCAAGCATTTTACTGTTATCAACCCGTTT	420
Db	436	GAGATGCTGGAGAGCTGACGCTCAGACCAAGCATTTTACTGTTATCAACCCGTTT	495
Qy	421	GGAGAAAGGACAAAGCAAGCGGATATGAAAGAAAGTGGCACCCTGTCACCTTA	480
Db	496	GGAGAAAGGACAAAGCAAGCGGATATGAAAGAAAGTGGCACCCTGTCACCTTA	555
Qy	481	GCCTCCATCACCCTGACATCATCGTTACTGAAACATGCTAATCAGGCGCAAGGAGACTG	540
Db	556	GCCTCCATCACCCTGACATCATCGTTACTGAAACATGCTAATCAGGCGCAAGGAGACTG	615
Qy	541	TATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGTTATGTT	600
Db	616	TATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGTTATGTT	675
Qy	601	AGCGAGGTGTCAGCGTCTGATTGGGAGGACGACAGAGAGCGCGGGTCCGACAGAAC	660
Db	676	AGCGAGGTGTCAGCGTCTGATTGGGAGGACGACAGAGAGCGCGGGTCCGACAGAAC	735
Qy	661	CACCCCGGGCTGTGCTGGTCCCCAGTAGCCTCCGATTGGAATCATTTCCCGCAGGTC	720
Db	736	CACCCCGGGCTGTGCTGGTCCCCAGTAGCCTCCGATTGGAATCATTTCCCGCAGGTC	795
Qy	721	ACGAGTCCGTGTGTTACTTCCACCGTGGCACCAGCAGCAGAGAACTCCGCGCTGCAT	780
Db	796	ACGAGTCCGTGTGTTACTTCCACCGTGGCACCAGCAGCAGAGAACTCCGCGCTGCAT	855

Qy	781	ATCGTTTGTGGGACTCGCTGGCCATGGATGTGTCTCTCAGTCCACCAACAGCACACTC	840
Db	856	ATCGTTTGTGGGACTCGCTGGCCATGGATGTGTCTCTCAGTCCACCAACAGCACACTC	915
Qy	841	CTTCGCTACTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGT	900
Db	916	CTTCGCTACTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGT	975
Qy	901	GAGAAGAAACGCTGTGGTGTGGGCTTGGCCAGATAGACATTTTCAGGTTTAAAGACCTT	960
Db	976	GAGAAGAAACGCTGTGGTGTGGGCTTGGCCAGATAGACATTTTCAGGTTTAAAGACCTT	1035
Qy	961	TCCCAACACTGCTATGAAAGGACAGATGCTCTCTCCCTGCAACACACACGTTGGATCT	1020
Db	1036	TCCCAACACTGCTATGAAAGGACAGATGCTCTCTCCCTGCAACACACACGTTGGATCT	1095
Qy	1021	CCAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGAGGCAAGCAAGCAGCAG	1080
Db	1096	CCAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGAGGCAAGCAAGCAGCAG	1155
Qy	1081	CTGGAGGAGGACGAGAAAGCACTGTATGTTTGAAGCTGCGAGGACGTTGGAGGAG	1140
Db	1156	CTGGAGGAGGACGAGAAAGCACTGTATGTTTGAAGCTGCGAGGACGTTGGAGGAG	1215
Qy	1141	TGGCAAGTCTGCTGTGGGAAAGTTTCTGGCCATCAATGCCCAAAATGCTCTGTGCTGT	1200
Db	1216	TGGCAAGTCTGCTGTGGGAAAGTTTCTGGCCATCAATGCCCAAAATGCTCTGTGCTGT	1275
Qy	1201	CGCGGAGCCCGGAGGCGCTCTCCCGGCTGCCCATCTTGGGAGACGGGTCTTCTGACCTC	1260
Db	1276	CGCGGAGCCCGGAGGCGCTCTCCCGGCTGCCCATCTTGGGAGACGGGTCTTCTGACCTC	1335
Qy	1261	ATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTTCTCATCAGCACCAAC	1320
Db	1336	ATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTTCTCATCAGCACCAAC	1395
Qy	1321	CAGCAGCAGCAGTTTGAATTTCACTTTTGTGAAAGTTTATCGCGTCAAGAAATTCAGTTT	1380
Db	1396	CAGCAGCAGCAGTTTGAATTTCACTTTTGTGAAAGTTTATCGCGTCAAGAAATTCAGTTT	1455
Qy	1381	ACGTGGAAGCAGATGAGGATGAGGACAGCAGCTCAAGAGGGGGGAGAGCGCTTT	1440
Db	1456	ACGTGGAAGCAGATGAGGATGAGGACAGCAGCTCAAGAGGGGGGAGAGCGCTTT	1515
Qy	1441	GGGCACATTTGACGACGACCCCTCTGCTGCTGTCACCGCTCTCAACAGCTCTCTGAAAC	1500
Db	1516	GGGCACATTTGACGACGACCCCTCTGCTGCTGTCACCGCTCTCAACAGCTCTCTGAAAC	1575
Qy	1501	TGCGACGGGAGGTCTGCGACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCGAGCTGGTT	1560
Db	1576	TGCGACGGGAGGTCTGCGACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCGAGCTGGTT	1635
Qy	1561	CGACTCTTTGCAAGAGAAATGGAAGAAATCCGAGGACGAGCTCAGCAGCTGA	1614
Db	1636	CGACTCTTTGCAAGAGAAATGGAAGAAATCCGAGGACGAGCTCAGCAGCTGA	1689

RESULT 2

ADJ96598 standard; DNA; 4429 BP.

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

XX AC ADJ96598;

gene; ds; kinase; human; SNP; single nucleotide polymorphism;
tyrosine protein kinase; serine/threonine protein kinase; PK; STK;
gene therapy; cancer; immune-related disease; cardiovascular disease;
brain; neuronal associated disease; metabolic; inflammatory disorder;
cytostatic; neuroprotective; immunomodulator; antiinflammatory;

KW	lipid kinase; KIAA1646.
XX	
OS	Homo sapiens.
OS	39.
XX	
PH	Key Location/Qualifiers
FT	variation replace(2391,g)
FT	/*tag= a
FT	/standard_name= "Single nucleotide polymorphism"
FT	variation replace(2577,g)
FT	/*tag= b
FT	/standard_name= "Single nucleotide polymorphism"
FT	variation replace(3020,c)
FT	/*tag= c
FT	/standard_name= "Single nucleotide polymorphism"
FT	variation replace(3769,c)
FT	/*tag= d
FT	/standard_name= "Single nucleotide polymorphism"
FT	variation replace(4272,g)
FT	/*tag= e
FT	/standard_name= "Single nucleotide polymorphism"
XX	
PX	WO2004006838-A2.
NN	
PD	22-JAN-2004.
XX	
XX	15-JUL-2003; 2003WO-US021730.
XX	
PR	15-JUL-2002; 2002US-0395632P.
XX	
PA	(SUGE-) SUGEN INC.
XX	
PI	Whyte D, Manning G, Caenepeel S;
XX	
WP	WI; 2004-122753/12.
DR	P-PSDB; ADJ96664.
XX	
PT	New nucleic acid molecule encoding a kinase polypeptide, useful for
PT	preparing a composition for treating diseases or disorders, e.g., cancer,
PT	or neurological, immunological or inflammatory disorders.
XX	
PS	Example 1; SEQ ID NO 55; 366pp; English.
XX	
CC	This invention relates to a novel isolated, enriched or purified nucleic
CC	acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC	to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
CC	as well as protein kinase-like enzymes. The present invention describes
CC	screening methods to identify agonists, antagonists and antibodies that
CC	can be used to modulate the activity or function of the mammalian kinase
CC	enzymes. As such, these compositions can be used for gene therapy
CC	purposes to treat diseases or disorders including cancer, immune-related
CC	diseases, cardiovascular disease, brain or neuronal associated disease,
CC	metabolic and inflammatory disorders. Accordingly, they exhibit
CC	cytostatic, neuroprotective, immunomodulator and antiinflammatory
CC	activities. This polynucleotide sequence is a human kinase DNA sequence
CC	of the invention.
XX	
SQ	Sequence 4429 BP; 1020 A; 1115 C; 1213 G; 1081 T; 0 U; 0 Other;
	Query Match 100.0%; Score 1614; DB 12; Length 4429;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 ATGGGGGCGACGGGGCGGCGCGCTGTGCTGCCTGTGTCGGAGCAGCACGCCG 60
Db	92 ATGGGGGCGACGGGGCGGCGCGCTGTGCTGCCTGTGTCGGAGCAGCACGCCG 151
QY	61 TGCGCCGCTGAGCCTTGAGAGCCCGCGGGCTCTGTGCTGTGTCGGAGCAGCACGCC 120
Db	152 TGCGCCGCTGAGCCTTGAGAGCCCGCGGGCTCTGTGCTGTGTCGGAGCAGCACGCC 211
QY	121 GGAGCGCGGCCCGCGCGCGAGTCCCTGTGCTGTATCTGAGATCATCGCGTT 180

Db	212	GGAGCCGGCGCCCCCGCGCGGATGCTGTCTGTGTGCTGTATCTGAGATCATCGCCGTT	271
Qy	181	GAGGAAACAGACGTTTACCGGAAACATCAAGGCAGTGGAAATCGCAGAAATCGAAAG	240
Db	272	GAGGAAACAGACGTTTACCGGAAACATCAAGGCAGTGGAAATCGCAGAAATCGAAAG	331
Qy	241	CTTTACGCTTTTACAGTTCACTGTGTTAAAGAGACACGACGGCACCGCTGGAAGTGGCG	300
Db	332	CTTTACGCTTTTACAGTTCACTGTGTTAAAGAGACACGACGGCACCGCTGGAAGTGGCG	391
Qy	301	CAGGTGACTTTCGGTGTCCAGAGGACAGCTGTGTCACTTGTGGCTGCAGACCTCGGG	360
Db	392	CAGGTGACTTTCGGTGTCCAGAGGACAGCTGTGTCACTTGTGGCTGCAGACCTCGGG	451
Qy	361	GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTTACTTGGTATTATCAACCCGTTT	420
Db	452	GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTTACTTGGTATTATCAACCCGTTT	511
Qy	421	GGAGGAAAGGACAAAGGCAAGCGGATATATGAAAGAAAGTGGCAACACTGTTTCAACCTTA	480
Db	512	GGAGGAAAGGACAAAGGCAAGCGGATATATGAAAGAAAGTGGCAACACTGTTTCAACCTTA	571
Qy	481	GCCTCCATCAGCACTGACATCATCGTTACTGAAATCATGCTTAATCAGGCCAAGGAGACTCTG	540
Db	572	GCCTCCATCAGCACTGACATCATCGTTACTGAAATCATGCTTAATCAGGCCAAGGAGACTCTG	631
Qy	541	TATGAGATTAACTAGACAAATACGACGGCATCGTCTGTGTCGCGCGAGATGGTATGTC	600
Db	632	TATGAGATTAACTAGACAAATACGACGGCATCGTCTGTGTCGCGCGAGATGGTATGTC	691
Qy	601	AGCGAGTGTGCACGGTCTGATTTGGGAGGACGACAGAGAGCGCGCGGGTGCACAGAAC	660
Db	692	AGCGAGTGTGCACGGTCTGATTTGGGAGGACGACAGAGAGCGCGCGGGTGCACAGAAC	751
Qy	661	CACCCCGGGTGTGCTGCCAGTAGCCTCCGGATTGGATTGAATCATTTCCGCGAGGGTCA	720
Db	752	CACCCCGGGTGTGCTGCCAGTAGCCTCCGGATTGGATTGAATCATTTCCGCGAGGGTCA	811
Qy	721	ACGAGCTGCTGTGTTACTCCACGTCGGGACACAGACGACGAGAAACCTCGGCGCTGCAT	780
Db	812	ACGAGCTGCTGTGTTACTCCACGTCGGGACACAGACGACGAGAAACCTCGGCGCTGCAT	871
Qy	781	ATCGTTTGTGGGACCTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAGCACACTC	840
Db	872	ATCGTTTGTGGGACCTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAGCACACTC	931
Qy	841	CTTCGCTACTCCGTCCTCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACGT	900
Db	932	CTTCGCTACTCCGTCCTCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACGT	991
Qy	901	GAGAGAAACGCTGTTGGGCTTGCAGATACGACTTTTCAGTTTAAAGACCTTCCTC	960
Db	992	GAGAGAAACGCTGTTGGGCTTGCAGATACGACTTTTCAGTTTAAAGACCTTCCTC	1051
Qy	961	TCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCGCAACACACGGTGGGATCT	1020
Db	1052	TCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCGCAACACACGGTGGGATCT	1111
Qy	1021	CCAAGGATAGGAAGCCCTCCGGGACAGATGCTTTGTTTTCAGGCAAAAGCAAGCAGAG	1080
Db	1112	CCAAGGATAGGAAGCCCTCCGGGACAGATGCTTTGTTTTCAGGCAAAAGCAAGCAGAG	1171
Qy	1081	CTGGAGGAGACAGAGAAAGCACTGTATGTTTGGAGCTCGGAGGACGTGGAGGAG	1140
Db	1172	CTGGAGGAGACAGAGAAAGCACTGTATGTTTGGAGCTCGGAGGACGTGGAGGAG	1231
Qy	1141	TGGCAAGTCTGCTGTGGAAAGTTTCTGGCCATCAATGCCCAAAACATGCTCTGTGCTGT	1200
Db	1232	TGGCAAGTCTGCTGTGGAAAGTTTCTGGCCATCAATGCCCAAAACATGCTCTGTGCTGT	1291
Qy	1201	CGCCGAGCCCCAGGGCCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	1260
Db	1292	CGCCGAGCCCCAGGGCCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	1351

```
Qy 1261 ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCATCAGGCACACCAAC 1320
D 1352 ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCATCAGGCACACCAAC 1411
Qy 1321 CAGCAGACCAAGTTGACTTCACATTTTGTGAAAGTTATCCGGTCAAGAAATCCAGTTT 1380
D 1412 CAGCAGGACCAAGTTGACTTCACATTTTGTGAAAGTTATCCGGTCAAGAAATCCAGTTT 1471
Qy 1381 ACCTCGAGACATGAGGATGAGACAGACCACTCAGAGGGGGGAGAGCGCTTT 1440
D 1472 ACCTCGAGACATGAGGATGAGACAGACCACTCAGAGGGGGGAGAGCGCTTT 1531
Qy 1441 GGGCACATTTGCAGCAGCACCCCTCTCTGCTGTCACCGCTCCAAACAGCTCTCTGGAAC 1500
D 1532 GGGCACATTTGCAGCAGCACCCCTCTCTGCTGTCACCGCTCTCAACAGCTCTGGAAC 1591
Qy 1501 TGCAGCGGGAGGTCCTGACACAGCCCTGCTGCTGTCACCGCTCAGAGTCCAGCTGGTT 1560
D 1592 TGCAGCGGGAGGTCCTGACACAGCCCTGCTGCTGTCACCGCTCAGAGTCCAGCTGGTT 1651
Qy 1561 CGACTCTTTGCAGCAGGAAATTCAGAGAAATCCGAGCCAGACTCACACAGCTGA 1614
D 1652 CGACTCTTTGCAGCAGGAAATTCAGAGAAATCCGAGCCAGACTCACACAGCTGA 1705
```

RESULT 3

ABX70921

ID ABX70921 standard; cDNA; 4432 BP.

XX

AC ABX70921;

XX

DT 05-MAR-2003 (first entry)

XX

DE Novel human cDNA sequence #146.

XX

KW Human; gene; ss; nervous system disorder; peripheral neuropathy;
KW Huntington's disease; amyotrophic lateral sclerosis; haemophilia;
KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
KW insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;
KW ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;
KW fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;
KW coagulation disorder; cancer; tumour; inflammatory disease; septic shock;
KW Crohn's disease; anaphylaxis; proliferation; chemotactic;
KW differentiation; stem cell growth factor; haematopoiesis; chemokinetic;
KW haemostatic; antiinflammatory; expressed sequence tag; EST.

XX Homo sapiens.

OS

XX WO200281731-A2.

PN

XX 17-OCT-2002.

PD

XX 29-JAN-2002; 2002WO-US001222.

XX

XX 30-JAN-2001; 2001US-00774528.

XX

XX (HYSE-) HYSEQ INC.

FA

XX (GOOD/) GOODRICH R W.

XX

XX Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX

XX Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;

XX

XX WPI; 2003-058563/05.

DR

XX Novel polypeptide useful for treating neurodegenerative diseases, myeloid

XX

XX or lymphoid cell disorders, bone disorders, mechanical and traumatic

XX

XX disorders, coagulation disorders, and inflammatory diseases.

XX

XX Claim 1; Page: 612pp; English.

PS

XX This invention relates to the cDNA sequences encoding an isolated novel

XX

CC human polypeptide. The protein encoded by the nucleic acid of the
CC invention is useful for treating central and peripheral nervous system
CC diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic
CC lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,
CC Alzheimer's disease); autoimmune disease (e.g. systemic lupus
CC erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)
CC ; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopaenia)
CC ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,
CC osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head
CC trauma); lung or liver fibrosis; reperfusion injury in various tissues;
CC bacterial, viral or fungal infections; allergic conditions such as
CC allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);
CC cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's
CC disease, anaphylaxis). The protein may be used to inhibit the growth,
CC infection or function of infectious agents such as bacteria, fungi,
CC viruses, or to effect bodily characteristics, biorhythms or circadian
CC cycles of rhythms. The protein may also have
CC proliferation/differentiation, stem cell growth factor, haematopoiesis
CC regulation, immune stimulating or suppressing, chemotactic/chemokinetic,
CC haemostatic and thrombolytic, receptor/ligand, and antiinflammatory
CC activities. The cDNA sequences of the invention are useful for expressing
CC recombinant protein for analysis. The present sequence represents a novel
CC human cDNA sequence of the invention, this sequence is an expressed
CC sequence tag (EST) and was identified using subtractive hybridisation
CC XX

SQ Sequence 4432 BP; 970 A; 1122 C; 1282 G; 1058 T; 0 U; 0 Other;

Query Match 99.9%; Score 1612.4; DB 8; Length 4432;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1613; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGGCGCAGCGGGCGCGGCGCGCGCGCTGCAATCCCTGTGGTGAAGCAGCAGCGC 60
D 1204 ATGGGGCGCAGCGGGCGCGGCGCGCGCGCTGCAATCCCTGTGGTGAAGCAGCAGCGC 1263

Qy 61 TGGCCCTGTAGCCTTGGAGCCCGCGCGGCTCTGCTGCGCTGTGGCGAGCCCGGGGCC 120
D 1264 TGGCCCTGTAGCCTTGGAGCCCGCGCGGCTCTGCTGCGCTGTGGCGAGCCCGGGGCC 1323

Qy 121 GGAGCCGGCCCCCGCGCGGAGTCCCTGCTGTGCTGTATCTGAGATCATCGCGTT 180
D 1324 GGAGCCGGCCCCCGCGCGGAGTCCCTGCTGTGCTGTATCTGAGATCATCGCGTT 1383

Qy 181 GAGGAAACAGACGTTTCACGGGAAACATCAAGCGCAGTGGAAAAATGGCAAAAATGGAAGA 240
D 1384 GAGGAAACAGACGTTTCACGGGAAACATCAAGCGCAGTGGAAAAATGGCAAAAATGGAAGA 1443

Qy 241 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACACGACCGCTGGAAGTGGCG 300
D 1444 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACACGACCGCTGGAAGTGGCG 1503

Qy 301 CAGGTGACTTCTGGTGTCCAGAGGAGCAGCTGTCTACTTGTGGCTGCAGACCTGCGG 360
D 1504 CAGGTGACTTCTGGTGTCCAGAGGAGCAGCTGTCTACTTGTGGCTGCAGACCTGCGG 1563

Qy 361 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCAATTTACTGGTATTTATCAACCCGTT 420
D 1564 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCAATTTACTGGTATTTATCAACCCGTT 1623

Qy 421 GGAGGAAAGGACAAAGGCAAGCGGATATATAAGAAAAAGTGGCACCACCTTACCTTA 480
D 1624 GGAGGAAAGGACAAAGGCAAGCGGATATATAAGAAAAAGTGGCACCACCTTACCTTA 1683

Qy 481 GCCTCCATCACCCTGACATCATCGTTTACTGAACTGTAATCAGGCAAGGAGACTCTG 540
D 1684 GCCTCCATCACCCTGACATCATCGTTTACTGAACTGTAATCAGGCAAGGAGACTCTG 1743

Qy 541 TATGAGATTAACATAGACAAATACGAGCGCATCGTCTGTGTCGGCGGAGATGGTATGTC 600
D 1744 TATGAGATTAACATAGACAAATACGAGCGCATCGTCTGTGTCGGCGGAGATGGTATGTC 1803

Qy 601 AGCAGGTGCTGACGGTCTGATTTGGGAGGACGAGAGCGCGCGGGGTGCAGCAGAAC 660

1804	Db	AGCAGAGTGTCTGACGGTCTGATTGGGAGAGACGACAGAGAGCGCGCGGGTTCGACACGAAAC	1863
661	Qy	CACCCCGGGTGTGCTGGTCCCGACGTAGCCTCCGGATTGGAAATCAATTCGCGCAGGGTCA	720
1864	Db	CACCCCGGGTGTGCTGGTCCCGACGTAGCCTCCGGATTGGAAATCAATTCGCGCAGGGTCA	1923
721	Qy	ACGCACTGCGTGTGTTACTCACCGTGGGACACAGCAGCAGGAAAACTTGGCGCGCTGCAT	780
1924	Db	ACGCACTGCGTGTGTTACTCACCGTGGGACACAGCAGCAGGAAAACTTGGCGCGCTGCAT	1983
781	Qy	ATCGTTGTTGGGGAAGCTCGCTGGCCATGGATGTGTCTTCAGTCCACCAACACAGCACACCTC	840
1984	Db	ATCGTTGTTGGGGAAGCTCGCTGGCCATGGATGTGTCTTCAGTCCACCAACACAGCACACCTC	2043
841	Qy	CTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGAACATCATCAAGGACAGT	900
2044	Db	CTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGAACATCATCAAGGACAGT	2103
901	Qy	GAGAAAGAAACGGTGGTGGGTCTTCCAGATACGACTTTTCAGSTTTTAAAGACCTTCCTC	960
2104	Db	GAGAAAGAAACGGTGGTGGGTCTTCCAGATACGACTTTTCAGSTTTTAAAGACCTTCCTC	2163
961	Qy	TCCCAACACTGCTATGAAGGAGCAGTGTCTTCTCCTGTCACAAACACACGGTGGGATCT	1020
2164	Db	TCCCAACACTGCTATGAAGGAGCAGTGTCTTCTCCTGTCACAAACACACGGTGGGATCT	2223
1021	Qy	CCAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAAAGCAAGCAGCAG	1080
2224	Db	CCAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAAAGCAAGCAGCAG	2283
1081	Qy	CTGGAGGAGGACAGAAAGCACTGTATGCTTTTGAAGCTGCGGAGGACGTGGAGGAG	1140
2284	Db	CTGGAGGAGGACAGAAAGCACTGTATGCTTTTGAAGCTGCGGAGGACGTGGAGGAG	2343
1141	Qy	TGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTGTGCTTGT	1200
2344	Db	TGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTGTGCTTGT	2403
1201	Qy	CGCGGAGCCCCCAGGGCCCTCTCCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC	1260
2404	Db	CGCGGAGCCCCCAGGGCCCTCTCCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC	2463
1261	Qy	ATCCTCATCGGAAATGCTCCAGGTTCAAATTTTCTGAGATTTCTCATCAGGCACACCAAC	1320
2464	Db	ATCCTCATCGGAAATGCTCCAGGTTCAAATTTTCTGAGATTTCTCATCAGGCACACCAAC	2523
1321	Qy	CAGCAGACCACTTTTGACTTTCACCTTTTGTGTAAGTTTATCCGCTCAAAGAAATTCAGT	1380
2524	Db	CAGCAGACCACTTTTGACTTTCACCTTTTGTGTAAGTTTATCCGCTCAAAGAAATTCAGT	2583
1381	Qy	ACGTGCAAGACATGGAGGATGAGGACAGCGACCTCAAGGAGGGGGGGAAGAGCGCTTT	1440
2584	Db	ACGTGCAAGACATGGAGGATGAGGACAGCGACCTCAAGGAGGGGGGGAAGAGCGCTTT	2643
1441	Qy	GGGCACATTTGACGACGCCACCCCTCTGCTGTGACACCGTCTCCACAGCTCCTTGAAC	1500
2644	Db	GGGCACATTTGACGACGCCACCCCTCTGCTGTGACACCGTCTCCACAGCTCCTTGAAC	2703
1501	Qy	TGCCAGCGGGAGGTCTTGCAACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCGCAGCTG	1560
2704	Db	TGCCAGCGGGAGGTCTTGCAACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCGCAGCTG	2763
1561	Qy	CGACTTTTGCACGAGGAATTGAAGAAATCCCGAAGCCCACTCACAACGCTGA	1614
2764	Db	CGACTTTTGCACGAGGAATTGAAGAAATCCCGAAGCCCACTCACAACGCTGA	2817

RESULT 4

ADP55247

ID ADP55247 standard: cDNA: 4445 BP.

XX

AC ADP55247;

XX

18-NOV-2004 (first entry)
Human PRO cDNA sequence SEQ ID NO:1223.
human; PRO; immune related disease; inflammatory immune response;
immune response stimulation; anti allergic; antianemic; antiarthritic;
antiasthmatic; antidiabetic; antinflammatory; antipsoriatic;
anti rheumatic; antithyroid; CNS; dermatological; gastrointestinal;
haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
neurotropic; neuroprotective; osteopathic; respiratory; vasotropic;
virucide; gene therapy; gene; ss.
Homo sapiens.
WO2004039956-A2.
13-MAY-2004.
28-OCT-2003; 2003WO-US034381.
29-OCT-2002; 2002US-0422472P.
(GETH) GENENTECH INC.
Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PW;
Wood WI, Wu TD;
WPI; 2004-376182/35.
P-PSDB; ADP55248.
New PRO polynucleotides and polypeptides, useful in diagnosing
and treating an immune related disease, e.g. systemic lupus
erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
stimulating an immune response.
Claim 2; SEQ ID NO 1223; 3009pp; English.
The present invention describes an isolated PRO nucleic acid (1). Also
described: (1) a vector comprising (1); (2) a host cell comprising the
vector of (1); (3) a process for producing a PRO polypeptides; (4) an
isolated PRO polypeptide; (5) a chimeric molecule comprising the
polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
antibody which specifically binds to a polypeptide of (4); (7) a
composition of matter comprising a polypeptide of (4), an agonist or
antagonist of the polypeptide or an antibody that binds to the
polypeptide in combination with a carrier; (8) an article of manufacture
comprising a container, a label on the container and a composition of
matter of (7); (9) a method of treating an immune related disease in a
mammal; (10) a method for determining the presence of a PRO polypeptide
in a sample suspected of having the polypeptide; (11) a method of
diagnosing an immune related disease or an inflammatory immune response
in mammal; (12) a method of identifying a compound that inhibits or
mimics the activity of or expression of a gene encoding a PRO polypeptide
; and (13) a method of stimulating the immune response in a mammal. The
PRO sequences have anti allergic, antianemic, antiarthritic,
antiasthmatic, antidiabetic, antinflammatory, antipsoriatic,
anti rheumatic, antithyroid, CNS, dermatological, gastrointestinal,
haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
neurotropic, neuroprotective, osteopathic, respiratory, vasotropic and
virucide activities, and can be used in gene therapy. The nucleic acid
(1) and the encoded polypeptides, compositions, kits and methods are
useful in diagnosing and treating an immune related disease and in
stimulating an immune response. The present sequence represents a human
PRO nucleotide sequence from the present invention.
Sequence 4445 BP; 1018 A; 1134 C; 1216 G; 1077 T; 0 U; 0 Other;
SQ

Query Match

Query Match
Best Local

Best local similarity 55.5%, Freq: NO: 0,
Matches 1613: Conservative 0: Mismatches

STOT 82732914

1
v

Y

QY 1 ATGGGGGCGACGGGGGCGGGAGCCGCTGCAATCCGCTGTGGGTGAAGCAGCAGCGC 60


```
Qy 961 TCCACCACTGCTATGAGGACAGTGTCTTCTCTCTCTGCAACAACACAGGTGGATCT 1020
Db 1036 TCCACCACTGCTATGAGGACAGTGTCTTCTCTCTCTGCAACAACACAGGTGGATCT 1095
Qy 1021 CCAAGGNTAGAACCCCTGCGGGCAGGATGCTTTGTTGACGCAAGCAGCAG 1080
Db 1096 CCAAGGNTAGAACCCCTGCGGGCAGGATGCTTTGTTGACGCAAGCAGCAG 1155
Qy 1081 CTGAGAGGAGCAGAAAGACACTGTATGTTTGAAGCTGCGGAGGAGCTGGAGAG 1140
Db 1156 CTGAGAGGAGCAGAAAGACACTGTATGTTTGAAGCTGCGGAGGAGCTGGAGAG 1215
Qy 1141 TGGCAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCAACATGTCTGTGCTGT 1200
Db 1216 TGGCAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCAACATGTCTGTGCTGT 1275
Qy 1201 CGCGGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGAGACGGTCTTCTGACCTC 1260
Db 1276 CGCGGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGAGACGGTCTTCTGACCTC 1335
Qy 1261 ATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCACCAAC 1320
Db 1336 ATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCACCAAC 1395
Qy 1321 CAGCAGGACCAAGTTGACTTCACTTTTGTGAAGTTTATCGCTCAAGAAATCCAGTTT 1380
Db 1396 CAGCAGGACCAAGTTGACTTCACTTTTGTGAAGTTTATCGCTCAAGAAATCCAGTTT 1455
Qy 1381 ACCTCGAAGCATGAGGATGAGGACAGCAGCTCAAGAGGGGGGGAAGCGCTTT 1440
Db 1456 ACCTCGAAGCATGAGGATGAGGACAGCAGCTCAAGAGGGGGGGAAGCGCTTT 1515
Qy 1441 GGGCACATTTTCAGCAGCACCCCTCTCTGCTGCTGCACTCCACAGCTCTGGAAC 1500
Db 1516 GGGCACATTTTCAGCAGCACCCCTCTCTGCTGCTGCACTCCACAGCTCTGGAAC 1575
Qy 1501 TGCAGCGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGCTGTT 1560
Db 1576 TGCAGCGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGCTGTT 1635
Qy 1561 CGACTCTTTGACAGGAAATTAAGAGAAATCGAAGCCAGACTCACAGCTGA 1614
Db 1636 CGACTCTTTGACAGGAAATTAAGAGAAATCGAAGCCAGACTCACAGCTGA 1689
```

RESULT 6

ABA96945
ID ABA96945 standard; cDNA; 4463 BP.

XX AC ABA96945;

XX DT 20-MAY-2002 (first entry)

XX DE Human ceramide kinase hCERK1-encoding cDNA.

XX KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;
neurological disease; inflammation; human immunodeficiency virus;
HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;
cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;
anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;
gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX CDS 124..1737

XX FT /*tag= a

XX FT /product= "Human ceramide kinase hCERK1"

XX PN WO200196575-A1.

XX PD 20-DEC-2001.

XX ,

PF 11-JUN-2001; 2001WO-JP004889.

XX PR 14-JUN-2000; 2000JP-00178039.

XX PA (SANY) SANKYO CO LTD.

XX PI Sugiura M, Kono K, Kohama T;

XX P-PSDB; AAM49115.

XX WPI; 2002-179513/23.

XX DR P-PSDB; AAM49115.

XX PT Human ceramide kinase gene and the enzyme encoded by it for screening

XX subences as drugs for neurological, inflammatory and other disorders.

XX PS Claim 5; Page 46-53; 61pp; Japanese.

XX CC This sequence represents cDNA encoding a human ceramide kinase designated

CC hCERK1. The invention relates to hCERK1, nucleic acids encoding it, the

CC expression vectors and host cells containing hCERK1 nucleic acids, the

CC recombinant production of hCERK1 and antibodies specific for hCERK1. The

CC invention also encompasses methods of isolating hCERK1 from samples, the

CC use of hCERK1 in drug screening, and the use of hCERK1 nucleic acid

CC sequences in gene therapy. hCERK1 mediates the ATP-dependent 1-

CC phosphorylation of ceramides and can be used to screen for therapeutic

CC and preventive agents for a wide range of disorders. Such disorders

CC include neurological disease, inflammation, human immunodeficiency virus

CC (HIV) infection, type 2 diabetes, obesity, sepsis, arteriosclerosis and

CC cancer

XX SQ Sequence 4463 BP; 1026 A; 1135 C; 1217 G; 1084 T; 0 U; 1 Other;

Query Match 99.6%; Score 1607.6; DB 6; Length 4463;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1610; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGGGCGCAGCGGGCGGCGGAGCCGCTGCAATCCGTCTGTGGTGAAGCAGCAGCGC 60

Db 124 ATGGGGCGCAGCGGGCGGCGGAGCCGCTGCAATCCGTCTGTGGTGAAGCAGCAGCGC 183

Qy 61 TGGCCGCTGAGCCCTGGAGCCGCGCGGCTCTGCTGGCTGGTGGCGAGCCGGGGGCC 120

Db 184 TGGCCGCTGAGCCCTGGAGCCGCGCGGCTCTGCTGGCTGGTGGCGAGCCGGGGGCC 243

Qy 121 GGAGCCGCGCCCGCGCGCGGATGCTGCTGTGCTGTATCTGAGATCATCGCGTT 180

Db 244 GGAGCCGCGCCCGCGCGCGGATGCTGCTGTGCTGTATCTGAGATCATCGCGTT 303

Qy 181 GAGGAAAACAGACGTTTACGGGAAAACATCAAGGCAAGTGGAAAAATGGCAAAAATGGAAGA 240

Db 304 GAGGAAAACAGACGTTTACGGGAAAACATCAAGGCAAGTGGAAAAATGGCAAAAATGGAAGA 363

Qy 241 CCTTACGCTTTTACAGTTTCTGTAAGAGAGACGACGCGCACCCCTGGAAGTGGGCG 300

Db 364 CCTTACGCTTTTACAGTTTCTGTAAGAGAGACGACGCGCACCCCTGGAAGTGGGCG 423

Qy 301 CAGGTGACTTTTCTGGTGTCCAGAGGACGAGCTGTGTCACTTGTGGTGCAGACCCCTGCGG 360

Db 424 CAGGTGACTTTTCTGGTGTCCAGAGGACGAGCTGTGTCACTTGTGGTGCAGACCCCTGCGG 483

Qy 361 GAGATGCTGGAGAAGCTGAGCTCCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTT 420

Db 484 GAGATGCTGGAGAAGCTGAGCTCCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTT 543

Qy 421 GGAGGAAAAGGACAAAGCGGATATATGAAAGAAAAGTGGCACCACCTGTTCACCTTA 480

Db 544 GGAGGAAAAGGACAAAGCGGATATATGAAAGAAAAGTGGCACCACCTGTTCACCTTA 603

Qy 481 GCCTCCATCACCCTGACATCATCGTTTACTGAAACATGCTATATCAGGCCAAGGAGACTCTG 540

Db 604 GCCTCCATCACCCTGACATCATCGTTTACTGAAACATGCTATATCAGGCCAAGGAGACTCTG 663

Qy 541 TATGAGATTAACTAGACAAATACGAGCGGATCGTCTGTGTCGGCGGAGATGGTATGTTTC 600

664	Db	TATGAGATTAA	CATAGACAA	AATACGACGG	CATCGTGTGTCGGCGGAGATGGTATGTTCT	723		
601	Qy	AGCGAGGTGCT	GCACGGTCT	GATTGGGAGGACG	CAGAGGAGCGCGGGGTCCACCAAGAAC	660		
724	Db	AGCGAGGTGCT	GCACGGTCT	GATTGGGAGGACG	CAGAGGAGCGCGGGTCCACCAAGAAC	783		
661	Qy	CACCCCGGGCT	GTGTGGTCC	CACAGTAGCCT	CCGGATTGGAATCATTTCCGCGAGGGTCA	720		
784	Db	CACCCCGGGCT	GTGTGGTCC	CACAGTAGCCT	CCGGATTGGAATCATTTCCGCGAGGGTCCA	843		
721	Qy	ACGGAATCGCT	GTGTACTCC	ACCGTGGGACAC	GAGCAGCAGAAAACTCGGGCGTCCAT	780		
844	Db	ACGGAATCGCT	GTGTACTCC	ACCGTGGGACAC	GAGCAGCAGAAAACTCGGGCGTCCAT	903		
781	Qy	ATCGTTGTTGG	GGGACTCGCT	GGCCATGGATGT	CTCAAGTCCACCAACAGCACATCT	840		
904	Db	ATCGTTGTTGG	GGGACTCGCT	GGCCATGGATGT	CTCAAGTCCACCAACAGCACATCT	963		
841	Qy	CTTCGCTACT	CCGTTCCCT	CTCGGCTTAC	GGGTCATCAGGGAGACATCATCAAGGACAGT	900		
964	Db	CTTCGCTACT	CCGTTCCCT	CTCGGCTTAC	GGGTCATCAGGGAGACATCATCAAGGACAGT	1023		
901	Qy	GAGAAGAAAC	GGTGGTGGT	CTTCCAGATAC	GACTTTTCAGGTTTAAAGACCTTCCCTC	960		
1024	Db	GAGAAGAAAC	GGTGGTGGT	CTTCCAGATAC	GACTTTTCAGGTTTAAAGACCTTCCCTC	1083		
961	Qy	TCCCAACACT	CTGTATGA	AGGACAGTGT	CTTCTCCTCGCATCAACACACGGTGGGATCT	1020		
1084	Db	TCCCAACACT	CTGTATGA	AGGACAGTGT	CTTCTCCTCGCATCAACACACGGTGGGATCT	1143		
1021	Qy	CCAAAGGATAG	AAAGCCCTG	CGCGGACAGGAT	GTCTTGTGTCAGGCAAAACGACAGCAGCAG	1080		
1144	Db	CCAAAGGATAG	AAAGCCCTG	CGCGGACAGGAT	GTCTTGTGTCAGGCAAAACGACAGCAGCAG	1203		
1081	Qy	CTGAGAGGAG	GACGAGAA	GAAGCACTGT	ATGGTTTGAAGCTGCGGAGGACGTGGAGGAG	1140		
1204	Db	CTGAGAGGAG	GACGAGAA	GAAGCACTGT	ATGGTTTGAAGCTGCGGAGGACGTGGAGGAG	1263		
1141	Qy	TGGCAAGT	CTGTGTGG	GAAGTTTCT	CGGCCATCAATGCCCAAAACATGCTCTGTGCTTGT	1200		
1264	Db	TGGCAAGT	CTGTGTGG	GAAGTTTCT	CGGCCATCAATGCCCAAAACATGCTCTGTGCTTGT	1323		
1201	Qy	CGCCGAGCCCC	CAGGGCCCT	CTCCC	CGGCTGCCACCTTGGGAGACGGGTCTTCTGACCTC	1260		
1324	Db	CGCCGAGCCCC	CAGGGCCCT	CTCCC	CGGCTGCCACCTTGGGAGACGGGTCTTCTGACCTC	1383		
1261	Qy	ATCCTCAT	CCGGAAATG	CTCTCAG	TTTCTGAGATTTCTCATCAGGCAACCAAC	1320		
1384	Db	ATCCTCAT	CCGGAAATG	CTCTCAG	TTTCTGAGATTTCTCATCAGGCAACCAAC	1443		
1321	Qy	CAGCAGAC	CAGTTTGACT	TTTTCAT	TTTGTGAGTTTATCGGTCAAGAAATTTCCAGTTT	1380		
1444	Db	CAGCAGAC	CAGTTTGACT	TTTTCAT	TTTGTGAGTTTATCGGTCAAGAAATTTCCAGTTT	1503		
1381	Qy	ACGTGCAAG	CACATPG	GAGATGAGG	ACAGCCTCAAGAGGGGGGGAAGAGCGCTTT	1440		
1504	Db	ACGTGCAAG	CACATPG	GAGATGAGG	ACAGCCTCAAGAGGGGGGGAAGAGCGCTTT	1563		
1441	Qy	GGGCACAT	TTTGCAG	CGCACCCCT	CTGCTGTGACCGTCTCCACAGCTCTCTGGAAC	1500		
1564	Db	GGGCACAT	TTTGCAG	CGCACCCCT	CTGCTGTGACCGTCTCCACAGCTCTCTGGAAC	1623		
1501	Qy	TGCGAC	CGGGAGGT	CTTGCA	CAGCCCTGSCCATCGAGGTCAAGTCCA	CTGCCAGCTGGTT	1560	
1624	Db	TGCGAT	GGGAGGT	CTTGCA	CAGCCCTGSCCATCGAGGTCAAGTCCA	CTGCCAGCTGGTT	1683	
1561	Qy	CGACTCT	TTTGCAC	GAGGAAT	TGAAGAAATCCG	AGCCAGACTCA	CACAGCTGA	1614
1684	Db	CGACTCT	TTTGCAC	GAGGAAT	TGAAGAAATCCG	AGCCAGACTCA	CACAGCTGA	1737

RESULT 7
ADN62844

ID	ADN62844	standard; DNA; 1740 BP.
XX		
AC	ADN62844;	
XX		
DT	01-JUL-2004	(first entry)
XX		
DE	Human NOV9a DNA.	
XX		
KW	ds; gene; human; NOVX; metabolic	
KW	infectious disease; anorexia; ca	
KW	neurodegenerative disorder; Alzh	
KW	immune disorder; haematopoietic	
XX	metabolic syndrome X; wasting di	
OS	Homo sapiens.	
XX		
PN	US2004038223-A1.	
XX		
PD	26-FEB-2004.	
XX		
PF	01-OCT-2002; 2002US-00262511.	
XX		
PR	02-OCT-2001; 2001US-0326483P.	
PR	05-OCT-2001; 2001US-0327435P.	
PR	05-OCT-2001; 2001US-0327449P.	
PR	09-OCT-2001; 2001US-0327917P.	
PR	09-OCT-2001; 2001US-0328029P.	
PR	09-OCT-2001; 2001US-0328044P.	
PR	09-OCT-2001; 2001US-0328056P.	
PR	12-OCT-2001; 2001US-0328849P.	
PR	15-OCT-2001; 2001US-0329414P.	
PR	17-OCT-2001; 2001US-0330142P.	
PR	18-OCT-2001; 2001US-0330309P.	
PR	22-OCT-2001; 2001US-0341058P.	
PR	24-OCT-2001; 2001US-0332668P.	
PR	24-OCT-2001; 2001US-0343623P.	
PR	29-OCT-2001; 2001US-0349575P.	
PR	01-NOV-2001; 2001US-0346357P.	
PR	17-APR-2002; 2002US-0373260P.	
PR	19-APR-2002; 2002US-0373815P.	
PR	19-APR-2002; 2002US-0373817P.	
PR	19-APR-2002; 2002US-0373826P.	
PR	19-APR-2002; 2002US-0373884P.	
PR	22-APR-2002; 2002US-0374977P.	
PR	16-MAY-2002; 2002US-0381038P.	
PR	16-MAY-2002; 2002US-0381039P.	
PR	16-MAY-2002; 2002US-0381042P.	
PR	18-MAY-2002; 2002US-0381642P.	
PR	28-MAY-2002; 2002US-0383656P.	
PR	29-MAY-2002; 2002US-0383831P.	
PR	25-JUN-2002; 2002US-0391335P.	
XX		
PA	(SMIT/ SMITHSON G.	
PA	(MILL/ MILLET I.	
PA	(PEYM/ PEYMAN J A.	
PA	(KEKU/ KEKUDA R.	
PA	(JUJU/ JU J.	
PA	(LILL/ LI L.	
PA	(GUOX/ GUO X.	
PA	(PATT/ PATTURAJAN M.	
PA	(SPYT/ SPYTEK K A.	
PA	(EDIN/ EDINGER S R.	
PA	(ELLE/ ELLERMAN K.	
PA	(MALV/ MALYANKAR U M.	
PA	(ORTT/ ORT T.	
PA	(GORM/ GORMAN L.	
PA	(ZERH/ ZERHUSSEN B D.	
PA	(ANDE/ ANDERSON D W.	
PA	(ZHON/ ZHONG M.	
PA	(CATT/ CATTERTON E.	
PA	(JIWW/ JI W.	
PA	(MILL/ MILLER C E.	
PA	(RAST/ RASTELLI L.	

ds; gene; human; NOVX; metabolic disorder; diabetes; obesity; infectious disease; anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immune disorder; haematopoietic disorder; dyslipidaemia; metabolic syndrome X; wasting disorder.

PA (STON/) STONE D J.
 PA (PENA/) PENA C E A.
 PA (SHEN/) SHENOY S G.
 PA (SHIM/) SHIMKETS R A.
 PA (ROTH/) ROTHENBERG M E.
 PA (LEACH/) LEACH M D.
 PA (AGEE/) AGEE M L.
 PA (BERG/) BERGHS C.
 PA (DIPI/) DIPIPO V A.
 PA (EISE/) EISEN A. A.
 PA (GANG/) GANGOLLI E A.
 PA (RIEG/) RIEGER D K.
 PA (SPAD/) SPADERNA S K.
 XX
 PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
 PI Ort T, Gorman L, Zernhusen BD, Anderson DW, Zhong M, Catterton E;
 PI Ji W, Miller CE, Rastelli L, Stone DV, Pena CE, Shenoy SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;
 XX
 DR WPI; 2004-213931/20.
 DR P-FSDB; ADN62845.
 XX
 PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.
 PT
 PS Claim 20; SEQ ID NO 39; 395pp; English.
 XX
 CC The invention relates to isolated NOVX polypeptides and polynucleotides.
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or
 CC treat a medical condition in human related to the aberrant expression and
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and
 CC polynucleotides may be used to treat disorders associated with decreased
 CC expression or activity of NOVX by supplementing the patient our
 CC production or to rectify mutations. Conversely, antisense NA molecules
 CC may be administered to down regulate expression of NOVX polypeptides by
 CC binding with the cells own genes and preventing their expression. NOVX
 CC polynucleotides and complementary sequences may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar sequences in samples, and so which patients may be in need of
 CC restorative therapy. NOVX polypeptides may also be used as antigens in
 CC the production of antibodies and in assays to identify modulators
 CC (agonists and antagonists) of the expression and activity of NOVX. The
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be
 CC used to modulate NOVX polynucleotide expression and activity of NOVX
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX
 CC polypeptides and polynucleotides may be used in this way to prevent,
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic
 CC disturbances associated with obesity, the metabolic syndrome X and
 CC wasting disorders associated with chronic diseases and various cancers.
 CC They may also be used as antibacterial agents. The present sequence
 CC represents DNA encoding a human NOVX protein.
 XX
 SQ Sequence 1740 BP; 388 A; 477 C; 530 G; 345 T; 0 U; 0 Other;

Query Match 98.3%; Score 1586.2; DB 12; Length 1740;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1610; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 ATGGGGGCGACGGGGGCGGAGCCGCTGCATCCGTGCTGGGTGAAGCAGCAGCGC 60
 |||||
 DB 76 ATGGGGGCGACGGGGGCGGAGCCGCTGCAATCCGTGCTGGGTGAAGCAGCAGCGC 135

QY 61 TGGCCGCTGAGCCTCGAGCCGCGGCGCTCTGCTCGCTGGTGGCGAGCCCGGGGCC 120
 |||||

DB 136 TGGCCGCTGAGCCTCGAGCCGCGGCGCTCTGCTCGCTGGTGGCGAGCCCGGGGCC 195
 |||||

QY 121 GGAGCCGGCGCCCGCGCGCGAGTGCCTGCTCTGTGCTGTATCTGAGATCATCGCCGTT 180
 |||||

DB 196 GGAGCCGGCGCCCGCGTGTGATGCTGCTGTGCTGTATCTGAGATCATCGCCGTT 255
 |||||
 QY 181 GAGGAACACAGCGTTCAACGGGAACATCAAGGCGAGTGGAAAAATGGCAGAAAAATGAAAAAG 240
 |||||
 DB 256 GAGGAACACAGCGTTCAACGGGAACATCAAGGCGAGTGGAAAAATGGCAGAAAAATGAAAAAG 315
 |||||
 QY 241 CCTTACCGCTTTTACAGTTCACTGTGTAAAGAGAGACACGACGGCACCCGCTGGAAGTGGCG 300
 |||||
 DB 316 CCTTACCGCTTTTACAGTTCACTGTGTAAAGAGAGACACGACGGCACCCGCTGGAAGTGGCG 375
 |||||
 QY 301 CAGGTGACTTTCTGCTGTCTCAGAGGAGCAGCTGTGTCACTTGTGGGTGCGAGACCTCGCG 360
 |||||
 DB 376 CAGGTGACTTTCTGCTGTCTCAGAGGAGCAGCTGTGTCACTTGTGGGTGCGAGACCTCGCG 435
 |||||
 QY 361 GAGATGCTGAGAGAGCTGAGCTCCAGACCAAGACATTTACTGTATTATTAACAACCCGTTT 420
 |||||
 DB 436 GAGATGCTGAGAGAGCTGAGCTCCAGACCAAGACATTTACTGTATTATTAACAACCCGTTT 495
 |||||
 QY 421 GGAGAAAAGGACAAAGCAAGCGGATATATGAAAGAAAAGTGGCACCACCTGTTTCACTTA 480
 |||||
 DB 496 GGAGAAAAGGACAAAGCAAGCGGATATATGAAAGAAAAGTGGCACCACCTGTTTCACTTA 555
 |||||
 QY 481 GCCTCCATCCACTGACATCATCTGTTACTGAACATGCTTAATCAGGCGCAAGGAGACTCTG 540
 |||||
 DB 556 GCCTCCATCCACTGACATCATCTGTTACTGAAATGCTTAATCAGGCGCAAGGAGACTCTG 615
 |||||
 QY 541 TATGAGATTAACTAGACAAATACAGCGCATCTGTGTGTCGGCGGAGAT-GGTATGTT 599
 |||||
 DB 616 TATGAGATTAACTAGACAAATACAGCGCAT-GTCTGTGTGCGCGGAGATCGGTATGTT 674
 |||||
 QY 600 CAGCGAGTGTGTGACCGGTCTGATTGGAGAGACGACAGAGGCGCCGGGTGCAACAGAA 659
 |||||
 DB 675 CAGCGAGTGTGTGACCGGTCTGATTGGAGAGACGACAGAGGCGCCGGGTGCAACAGAA 734
 |||||
 QY 660 CCACCCCGGGCTGTGTGTCCTCCAGTAGCCTCCGATGGGAATCATTTCCCGCAGGGTC 719
 |||||
 DB 735 CCACCCCGGGCTGTGTGTCCTCCAGTAGCCTCCGATGGGAATCATTTCCCGCAGGGTC 794
 |||||
 QY 720 AACGGACTGGGTGTTTACTCCACCGTGGGACACGACGACGACAGAAAACCTCGGCGCTGCA 779
 |||||
 DB 795 AACGGACTGGGTGTTTACTCCACCGTGGGACACGACGACGACAGAAAACCTCGGCGCTGCA 854
 |||||
 QY 780 TATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACACACACACACACT 839
 |||||
 DB 855 TATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACACACACACACACT 914
 |||||
 QY 840 CTTTCGCTACTCCGTCCTGTCCTGGGTACGGCTTCTACGGGACATCATCAAGGACAG 899
 |||||
 DB 915 CTTTCGCTACTCCGTCCTGTCCTGGGTACGGCTTCTACGGGACATCATCAAGGACAG 974
 |||||
 QY 900 TGAGAAGAAACGGTGGTGGTCTTGCAGATACGACTTTTCAGGTTTAAAGACTTCTCT 959
 |||||
 DB 975 TGAGAAGAAACGGTGGTGGTCTTGCAGATACGACTTTTCAGGTTTAAAGACTTCTCT 1034
 |||||
 QY 960 CTCCACACACTGCTATGAAGGACAGTGTCTCTCTCCCTGACACACACACCGTGGGATC 1019
 |||||
 DB 1035 CTCCACACACTGCTATGAAGGACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1094
 |||||
 QY 1020 TCCAAGGATAGGAAGCCCTGCGCGGACAGGATGCTTTGTTGACGGCAAGCAAGCAGCA 1079
 |||||
 DB 1095 TCCAAGGATAGGAAGCCCTGCGCGGACAGGATGCTTTGTTGACGGCAAGCAAGCAGCA 1154
 |||||
 QY 1080 GCTCGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTCGGAGGACGCTGGAGGA 1139
 |||||
 DB 1155 GCTCGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTCGGAGGACGCTGGAGGA 1214
 |||||
 QY 1140 GTGCAAGTCTGTGTGGAGATTTCTGGCCATCAATGCCCCAACAATGTCTCTGTGCTG 1199
 |||||
 DB 1215 GTGCAAGTCTGTGTGGAGATTTCTGGCCATCAATGCCCCAACAATGTCTCTGTGCTG 1274
 |||||
 QY 1200 TCGCGGAGGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGCTCTCTGACCT 1259
 |||||

Db 1275 TCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGGGGTCTTCTGACT 1334
Qy 1260 CATCTCATCCGGAATGCTCAGGTTCAATTTCTGAGATTCTCATCAGGCACACAA 1319
Db 1335 CATCTCATCCGGAATGCTCAGGTTCAATTTCTGAGATTCTCATCAGGCACACAA 1394
Qy 1320 CCAGCAGGACCATGTTGACTTCACTTTTCTGAAAGTTTATCGCTCAAGAAATTCAGTT 1379
Db 1395 CCAGCAGGACCATGTTGACTTCACTTTTCTGAAAGTTTATCGCTCAAGAAATTCAGTT 1454
Qy 1380 TAGCTCGAAGCATGAGGATGAGGACAGCAGCTTCAAGGAGGGGGGGAAGCGCTT 1439
Db 1455 TAGCTCGAAGCATGAGGATGAGGACAGCAGCTTCAAGGAGGGGGGGAAGCGCTT 1514
Qy 1440 TGGGCACATTTGAGCAGCAGCCCTCTCTGCTGCTGACCCGCTTCAACAGCTCTCTGGA 1499
Db 1515 TGGGCACATTTGAGCAGCAGCCCTCTCTGCTGCTGACCCGCTTCAACAGCTCTCTGGA 1574
Qy 1500 CTGCGACGGGAGGTCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGTGT 1559
Db 1575 CTGCGACGGGAGGTCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGTGT 1634
Qy 1560 TCGACTCTTTGACGAGGAAATTGAAGAGAAATCCGAAGCCAGACTCACACAGCTGA 1614
Db 1635 TCGACTCTTTGACGAGGAAATTGAAGAGAAATCCGAAGCCAGACTCACACAGCTGA 1689

RESULT 8

AAA50510
ID AAA50510 standard; cDNA; 4231 BP.
XX
AC AAA50510;
XX
DT 05-DEC-2000 (first entry)
XX
DE Human sphingosine kinase C cDNA.
XX
KW Sphingosine kinase C; SKC; human; drug screening; infection;
KW antiinflammatory; antiallergic; anticancer; inflammation; allergy;
KW cancer; therapy; diagnosis; ds.
XX

OS Homo sapiens.

Key Location/Qualifiers
CDS 71..1453
/tag= a

WO200052173-A2.

PD 08-SEP-2000.

PF 02-MAR-2000; 2000WO-CA000223.

PR 02-MAR-1999; 99US-0122516P.

XX (ALIX) NPS ALLELIX CORP.

XX Munroe D, Gupta A, Falzone GR;

XX WPI: 2000-572185/53.

DR P-PSDB; AAY96059.

XX New human sphingosine kinase A, B and C polynucleotides and polypeptides
XX useful in e.g. chromosome and gene mapping, and detecting inflammation or
XX disease associated with abnormal levels of sphingosine kinase expression.

XX Disclosure; Fig 7; 81pp; English.

XX The present sequence is that of an isolated polynucleotide encoding human
XX sphingosine kinase C (SKC, see AAY96059), an enzyme that phosphorylates
XX sphingosine to form sphingosine 1-phosphate. The polynucleotide was
XX isolated from an HeLa cDNA library by PCR amplification. The invention
XX provides polynucleotides (see AAA50508-10) and polypeptides (see AAY96057

CC -59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC.
CC The polynucleotides may be used as hybridization probes, in the
CC construction of PCR primers for chromosome and gene mapping, in the
CC recombinant production of SKA, SKB and SKC, and in the generation of
CC antisense DNA or RNA. They can be used to detect inflammation or disease
CC associated with abnormal levels of SK expression, or to detect
CC differences in gene sequence between normal and carrier or affected
CC individuals. Host cells expressing SK can be used in drug screening.
CC Human SK specific antibodies, inhibitors, ligands or their analogues are
CC useful as bioactive agents to treat inflammation or disease including
CC viral, bacterial or fungal infections, allergic responses, mechanical
CC injury associated with trauma, hereditary diseases, lymphoma or
CC carcinoma, and other conditions with activate the genes of kidney, lung,
CC heart, lymphoid or tissues of the nervous system
XX

SQ Sequence 4231 BP; 1022 A; 1021 C; 1086 G; 1067 T; 0 U; 35 Other;

Query Match 89.7%; Score 1448.2; DB 3; Length 4231;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 162 ATCTGAGATCATCCGTTGAGGAAACAGAGCTTTCACGGGAAACATCAAGGCAGTGGAAA 221
Db 1 ATCTGAGATCATCCGTTGAGGAAACAGAGCTTTCACGGGAAACATCAAGGCAGTGGAAA 60
Qy 222 ATGCGAGAAATGGAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGAG 281
Db 61 ATGCGAGAAATGGAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGAG 120
Qy 282 GCACGCTGGAAGTGGCGCAGGTGACCTTCTGGTGTCAGAGGAGCAGCTGTGTACATT 341
Db 121 GCACGCTGGAAGTGGCGCAGGTGACCTTCTGGTGTCAGAGGAGCAGCTGTGTACATT 180
Qy 342 GTGGCTGCAGACCCCTGCGGAGATGCTGGGAAAGCTGACGTCACGACCAAGCAATTTACT 401
Db 181 GTGGCTGCAGACCCCTGCGGAGATGCTGGGAAAGCTGACGTCACGACCAAGCAATTTACT 240
Qy 402 GGTATTTATCAACCGTTTGGAGGAAAGGACAAGGCAAGCGGATATATGAAAGAAAGT 461
Db 241 GGTATTTATCAACCGTTTGGAGGAAAGGACAAGGCAAGCGGATATATGAAAGAAAGT 300
Qy 462 GGCACCACTGTTACCTTAGCTCCATCACCACATGACATCATCGTTACTGAACATGCTAA 521
Db 301 GGCACCACTGTTACCTTAGCTCCATCACCACATGACATCATCGTTACTGAACATGCTAA 360
Qy 522 TCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCGATCGTCTGTGT 581
Db 361 TCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCGATCGTCTGTGT 420
Qy 582 CGCGGAGATGGTATGTTTACGAGGAGTCTGCACGGTCTGATTGGGAGGACGAGAGGAG 641
Db 421 CGCGGAGATGGTATGTTTACGAGGAGTCTGCACGGTCTGATTGGGAGGACGAGAGGAG 480
Qy 642 CGCGGGGTGCACAGAACACCCCGGGCTGTCTGTGTCGCCAGTAGCCCTCCGAGTTGG 701
Db 481 CGCGGGGTGCACAGAACACCCCGGGCTGTCTGTGTCGCCAGTAGCCCTCCGAGTTGG 540
Qy 702 AATCATTTCCCGCAGGGTCAACGAGCTGCTGTGTTTACTTCCACCGTGGGCACGAGCAGCG 761
Db 541 AATCATTTCCCGCAGGGTCAACGAGCTGCTGTGTTTACTTCCACCGTGGGCACGAGCAGCG 600
Qy 762 AGAAACCTCGCGCTGCATATCGTTGTTGGGGAATCGCTGGCCATGGATGTCTCTAGT 821
Db 601 AGAAACCTCGCGCTGCATATCGTTGTTGGGGAATCGCTGGCCATGGATGTCTCTAGT 660
Qy 822 CCACCAACAGCAGCACTCTTCCGCTACTCCGCTCTCCCTGCTGGCTACGGCTTCTACGG 881
Db 661 CCACCAACAGCAGCACTCTTCCGCTACTCCGCTCTCCCTGCTGGCTACGGCTTCTACGG 720
Qy 882 GGACATCATCAAGGACAGTGAAGAAACGGTGGTGTGGGTCTTCCAGATACGACTTTTC 941
Db 721 GGACATCATCAAGGACAGTGAAGAAACGGTGGTGTGGGTCTTCCAGATACGACTTTTC 780


```
Db 481 ATCGTCTGTGTCGGCGGAGATGGTATGTTTCAGGAGGTGCTGCACGGTCTGATTGGGAGG 540
QY 631 AGCAGAGAGGCGCGGGGTGCACAGAAACCAACCCCGGGCTGTGTCTGCTCCCAAGTAGC 690
Db 541 ACCGAGAGAGCGCGGGGTGCACAGAAACCAACCCCGGGCTGTGTCTGCTCCCAAGTAGC 600
QY 691 CTCGGATTGGAAATCATTTCCCGAGGTCACCGAGGTCAACGGACTGGTGTACTTCCACCGTGGC 750
Db 601 CTCGGATTGGAAATCATTTCCCGAGGTCACCGAGGTCAACGGACTGGTGTACTTCCACCGTGGC 660
QY 751 ACCAGCGACGCAAAACCTTCGGCGCTGCATATCTGTTTGGGGACTCGCTGGCCATGGAT 810
Db 661 ACCAGCGACGCAAAACCTTCGGCGCTGCATATCTGTTTGGGGACTCGCTGGCCATGGAT 720
QY 811 GTGTCTCTAGTCACCAACAAGCACATCTCTTGTGCTACTCCGTGTCTCTGTGGGCTAC 870
Db 721 GTGTCTCTAGTCACCAACAAGCACATCTCTTGTGCTACTCCGTGTCTCTGTGGGCTAC 780
QY 871 GGTCTTCTACGGGACATCATCAGGACAGTGAAGAACGGTGGTGGTCTTCCCA 930
Db 781 GGTCTTCTACGGGACATCATCAGGACAGTGAAGAACGGTGGTGGTCTTCCCA 840
QY 931 TAGGACTTTTCAAGTTTAAAGACCTTCTCTCCCACTGTATGAAGGACAGTGTCC 990
Db 841 TAGGACTTTTCAAGTTTAAAGACCTTCTCTCCCACTGTATGAAGGACAGTGTCC 900
QY 991 TTCTCTCTCTGCAACAACACGCTGGGATCTCAAGGATAGAAAGCCCTGCGGACAGA 1050
Db 901 TTCTCTCTCTGCAACAACACGCTGGGATCTCAAGGATAGAAAGCCCTGCGGACAGA 960
QY 1051 TCGTTTGTGTCAGGCAAGACAGCAGCTGTGGAGGAGGACAGAAAGACATGTAT 1110
Db 961 TCGTTTGTGTCAGGCAAGACAGCAGCTGTGGAGGAGGACAGAAAGACATGTAT 1020
QY 1111 GGTTTGGAAGCTCGGAGACCTGGAGAGTGGCAAGTCTGTGGGAAGTTTCTGGCC 1170
Db 1021 GGTTTGGAAGCTCGGAGACCTGGAGAGTGGCAAGTCTGTGGGAAGTTTCTGGCC 1080
QY 1171 ATCAATGCCAACAACATGCTGTGTGTCGCGGAGCCCGAGGGGCTCTCCCGGT 1230
Db 1081 ATCAATGCCAACAACATGCTGTGTGTCGCGGAGCCCGAGGGGCTCTCCCGGT 1140
QY 1231 GCCACTTGGGAGAGCGGTCTTGTGACCTCATCTCATCCGGAATGCTCAGGTTCAAT 1290
Db 1141 GCCCACTTGGGAGAGCGGTCTTGTGACCTCATCTCATCCGGAATGCTCAGGTTCAAT 1200
QY 1291 TTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCTTTGACTTTCATTTTGT 1350
Db 1201 TTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCTTTGACTTTCATTTTGT 1260
QY 1351 GAAGTTTATCGCGTCAAGAAATTCAGTTTACGTTCGAAGCATGAGGATGAGGACAGC 1410
Db 1261 GAAGTTTATCGCGTCAAGAAATTCAGTTTACGTTCGAAGCATGAGGATGAGGACAGC 1320
QY 1411 GACCTCAAGGAGGGGGAGAGCGCTTTGGGCAATTTGACAGAGCACCCTCTCTGC 1470
Db 1321 GACCTCAAGGAGGGGGAGAGCGCTTTGGGCAATTTGACAGAGCACCCTCTCTGC 1380
QY 1471 TCGTCACCGCTCTCAACAGCTCTTGGAACTCGGACGGGAGGTCTGACAGCCCTGCC 1530
Db 1381 TCGTCACCGCTCTCAACAGCTCTTGGAACTCGGACGGGAGGTCTGACAGCCCTGCC 1440
QY 1531 ATCGAGGTTCAGAGTCCATGCGAGCTGGTTGACTCTTTGACAGGAGAAATGAAGAGAT 1590
Db 1441 ATCGAGGTTCAGAGTCCATGCGAGCTGGTTGACTCTTTGACAGGAGAAATGAAGAGAT 1500
QY 1591 CCGAAGCCAGACTCACACAGCTGA 1614
Db 1501 CCGAAGCCAGACTCACACAGCTGA 1524
```

RESULT 10
ACC70838

```
ID ACC70838 standard; DNA; 3975 BP.
XX
AC ACC70838;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human Sphingosine kinase 4 coding sequence.
XX
KW Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;
platelet transfusion; platelet stabiliser; gene; da.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1251
FT /tag= a
FT /product= "Sphingosine kinase 4"
XX
PN WO2003031627-A1.
XX
PD 17-APR-2003.
XX
PF 28-SEP-2001; 2001WO-JP008537.
XX
PR 28-SEP-2001; 2001WO-JP008537.
XX
PA (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.
PA (CHBI-) CHEM BIOLOGY INST.
XX
PI Igarashi Y, Kihara A;
XX
DR WPI; 2003-354917/33.
DR P-PSDB; ABR56301.
XX
PT Platelet derived polypeptides with sphingosine kinase activity for
treatment of sphingosine related disorders.
XX
PS Claim 4; Page 30-31; 39pp; Japanese.
XX
CC The present sequence is the coding sequence for human sphingosine kinase
4 (SPHK4). The kinase can be used for the diagnosis and treatment of
sphingosine related disorders. The kinase can also be potentially used
for controlling toxicity of platelet transfusion and as a platelet
stabiliser
XX
SQ Sequence 3975 BP; 943 A; 978 C; 1042 G; 1012 T; 0 U; 0 Other;

Query Match 77.5%; Score 1251; DB 8; Length 3975;
Best Local Similarity 100.0%; Pred. No. 2.8e-314;
Matches 1251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 ATGCTGGAGAAAGCTGACGTCCAGACCAGCAATTTACTGGTATTTATCAACCCGTTTGA 423
Db 1 ATGCTGGAGAAAGCTGACGTCCAGACCAGCAATTTACTGGTATTTATCAACCCGTTTGA 60
QY 424 GGAAGAGCAAGGCAAGCGGATATATGAAAGAAAAGTGGACCACTGTTTACCTTAGCC 483
Db 61 GGAAGAGCAAGGCAAGCGGATATATGAAAGAAAAGTGGACCACTGTTTACCTTAGCC 120
QY 484 TCCATCACCACCTGACATCATCTGTTACTGAAATGCTATATCAGGCCAAGAGACTCTGTAT 543
Db 121 TCCATCACCACCTGACATCATCTGTTACTGAAATGCTATATCAGGCCAAGAGACTCTGTAT 180
QY 544 GAGATTAACTAGACAAATACGCGCATCTGTGTGCGCGGAGATGGTATGTTTCAGC 603
Db 181 GAGATTAACTAGACAAATACGCGCATCTGTGTGCGCGGAGATGGTATGTTTCAGC 240
QY 604 GAGGTGCTGACGGTCTGATTTGGGAGGACGACAGGAGCGCGGGGTGCAGACAGAACAC 663
Db 241 GAGGTGCTGACGGTCTGATTTGGGAGGACGACAGGAGCGCGGGGTGCAGACAGAACAC 300
QY 664 CCCGGGCTGTGTGGTCCCCCAGTAGCCTCCGATTTGAATCATTCGCCGAGGGTCAACG 723
|||||
```

Db 301 CCCCGGCTGTGCTGCTGCCAGTAGCTCCGGAATGGAATCATTCGCCGAGGCTCAACG 360
Qy 724 GACTCGCTGTGTACTCCACCGTGGGACACGACGACGAGAACTCTGGCGCTGCATATC 783
Db 361 GACTCGCTGTGTACTCCACCGTGGGACACGACGACGAGAACTCTGGCGCTGCATATC 420
Qy 784 GTTGTGGGACTCGCTGGCCATGATGTCTCTAGTCCACCAACAGACACTCCTT 843
Db 421 GTTGTGGGACTCGCTGGCCATGATGTCTCTAGTCCACCAACAGACACTCCTT 480
Qy 844 CGCTACTCCGTCTCCCTGCTGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTAG 903
Db 481 CGCTACTCCGTCTCCCTGCTGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTAG 540
Qy 904 AAGAAACGGTGTGGGTCTTGCAGATACGACTTTTTCAGGTTTTAAAGACTTCTCTCC 963
Db 541 AAGAAACGGTGTGGGTCTTGCAGATACGACTTTTTCAGGTTTTAAAGACTTCTCTCC 600
Qy 964 CACCACTGCTATGAGGACAGTGTCTCTCTCCATGCAACACACAGGTGGGATCTCCA 1023
Db 601 CACCACTGCTATGAGGACAGTGTCTCTCTCCATGCAACACACAGGTGGGATCTCCA 660
Qy 1024 AGGATAGGAAGCCCTGCCGGGACGAGTCTTGTTCAGGCAAGCAAGCAGCAGTG 1083
Db 661 AGGATAGGAAGCCCTGCCGGGACGAGTCTTGTTCAGGCAAGCAAGCAGCAGTG 720
Qy 1084 GAGGAGGACAGAAAGACACTGTATGTTTGAAGCTCGGAGGAGCTGGAGAGTGG 1143
Db 721 GAGGAGGACAGAAAGACACTGTATGTTTGAAGCTCGGAGGAGCTGGAGAGTGG 780
Qy 1144 CAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTGTGCTGTGCG 1203
Db 781 CAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTGTGCTGTGCG 840
Qy 1204 CGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTCGACTCATC 1263
Db 841 CGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTCGACTCATC 900
Qy 1264 CTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGACACCAACCAG 1323
Db 901 CTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGACACCAACCAG 960
Qy 1324 CAGGACCAAGTTGACTTCACATTTTGTGAAGTTTATCCGCTCAAGAAATTCAGTTTACG 1383
Db 961 CAGGACCAAGTTGACTTCACATTTTGTGAAGTTTATCCGCTCAAGAAATTCAGTTTACG 1020
Qy 1384 TCGAAGCACATGAGGATGAGACAGCACCTCAAGGAGGGGGGAGAGCGCTTTGGG 1443
Db 1021 TCGAAGCACATGAGGATGAGGACAGCACCTCAAGGAGGGGGGAGAGCGCTTTGGG 1080
Qy 1444 CACATTTGCAGCAGCACCCCTCTCTGCTGCTGCAACCGCTCTCAACAGCTCTCTGGAACGTC 1503
Db 1081 CACATTTGCAGCAGCACCCCTCTCTGCTGCTGCAACCGCTCTCAACAGCTCTCTGGAACGTC 1140
Qy 1504 GACGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCAGAGTCCATGCTCCAGCTGGTTGGA 1563
Db 1141 GACGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCAGAGTCCATGCTCCAGCTGGTTGGA 1200
Qy 1564 CTCCTTGCAGGAGATTGAGAGNATCCGAGCCAGACTCACACAGCTGA 1614
Db 1201 CTCCTTGCAGGAGATTGAGAGNATCCGAGCCAGACTCACACAGCTGA 1251

RESULT 11
ID ADS10370 standard; DNA; 4702 BP.
XX
AC ADS10370;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human therapeutic DNA - SEQ ID 607.
XX

KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulneryary;
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
XX aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
OS Homo sapiens.
XX
PN WO2004080148-A2.
PD 23-SEP-2004.
XX
PF 30-SEP-2003; 2003WO-US030720.
XX
PR 02-OCT-2002; 2002US-0416186P.
XX (NUVE-) NUVELO INC.
XX
PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AU, Weng G, Zhou P;
XX WPI; 2004-668857/65.
XX P-PSDB; ADS11054.
XX
PT New polynucleotide, useful in preparing a composition for diagnosing or
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT aplastic anemia or cancer for promoting wound healing.
XX
PS Claim 1; SEQ ID NO 607; 718pp; English.
XX
CC The invention relates to a novel isolated polynucleotide and the encoded
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during gene therapy
CC procedures. The current sequence is that of a human therapeutic DNA of
CC the invention. The current sequence is not shown explicitly within the
CC specification but can be accessed from the WIPO web-site.
XX
SQ Sequence 4702 BP; 970 A; 1226 C; 1364 G; 1118 T; 0 U; 24 Other;

Query Match 63.0%; Score 1017.6; DB 13; Length 4702;
Best Local Similarity 81.4%; Pred. No. 1e-253;
Matches 1261; Conservative 0; Mismatches 214; Indels 75; Gaps 4;
Qy 140 CGGATGCTGCTGTGCTGCTGTATCTGAGATCATCGCGTTGAGGAAACAGCCTTCACG 199
Db 1394 CAGATGCTGCTGTGCTGCTGTATCTGAGATCATCGCGTTGAGGAAACAGCCTTCACG 1453
Qy 200 GGAAACATCAAGGACGTGGAAATGGCAGAAATGGAAAGCCTTTACGCTTTTACAGTTC 259
Db 1454 GGAAACATCAAGGACGTGGAAATGGCAGAAATGGAAAGCCTTTACGCTTTTACAGTTC 1513
Qy 260 ACTGTGTAAGAGAGACGACGCGCACCTGGAAATGGGCGCAGGTGACTTCTGCTGTC 319
Db 1514 ACTGTGTAAGAGAGACGACGCGCACCTGGAAATGGGCGCAGGTGACTTCTGCTGTC 1573
Qy 320 CAGAGGACGAGCTGTGTCACCTGTGGCTGCAAGCCCTGCGGAGATGCTGGAGAGCTGA 379
Db 1574 CAGAGGACGAGCTGTGTCACCTGTGGCTGCAAGCCCTGCGGAGATGCTGGAGAGCTGA 1633
Qy 380 CGTCCAGAC-----CAAAGCATTTTACTGCTATTTTATCAACCCGCTTTGGAGAA 427
Db 1634 TTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACAGACAAATACG 1693
Qy 428 AAGGACAAAGGACGAGATATGAAAGAAAGTGGCACCACCTGTTACCTTAGCTCCCA 487
Db 1694 ACGGCATCGTCTGTGTCGGCGGAGATGTTGTTTTCAGCGAGGTGCTGCACCGTCTGATG 1753
Qy 488 TCACCACTGACATCATCGTTTACTCAACATGCTAATCTAGGCCCAAGGAGACTCTGTATGAGA 547
Db 1754 GGAGGACGACAGGAGCGCGGGTCCAGCAGAACCCCGCGGCTGTGTGCTGCCCA 1813

QY 548 TTAACAT-----AGACAAATACGACGATCGTCTGTCTGCGGAGATG----- 592
Db 1814 GTAGCTCCGGATTGAATCATTCGCCAGAGCTTTGCAATGACCTGGCGGAGGAGG 1873
QY 593 -GTATGTTGAGGAGGTGTGACGGTCTGATTTGGAGGACGACGAGGAGCCCGGGTTC 651
Db 1874 TGTCTGTCTCTCTGCGCCCTGTCTGTGCGCCGAGGGTGGCGCATGGTGCACTTTCA 1933
QY 652 GACCAGAACCCCGGCTGTGCTGCTCCAGTAGCTCGGATCGGAATCATTCCTCC 711
Db 1934 CTCTGTGACAGTCTGGGATGTGAGACCGCAGTCATCCCATTTATGGATGAAGACA 1993
QY 712 GCA-----GGGTCAACGG 724
Db 1994 GGAGACTGGGAGCATGTGGSCCCCGGTGAGAACCTGTGGTGTGGAGGGTCAACGG 2053
QY 725 ACTCGTGTGTTACTCCACCGTGGGACGACGAGGAGCAAAACCTCGGCGCTGCATATCG 784
Db 2054 ACTCGTGTGTTACTCCACCGTGGGACGACGAGGAGCAAAACCTCGGCGCTGCATATCG 2113
QY 785 TTGTTGGGACTCGCTGGGACATGATGTCTCTAGTCCACCAACAGCACACTCTCTTC 844
Db 2114 TTGTTGGGACTCGCTGGGACATGATGTCTCTAGTCCACCAACAGCACACTCTCTTC 2173
QY 845 GCTACTCTGCTGCTGCTGGGCTACGGCTTCTAGGGGACATCATCAAGGACAGTGAGA 904
Db 2174 GCTACTCTGCTGCTGCTGGGCTACGGCTTCTAGGGGACATCATCAAGGACAGTGAGA 2233
QY 905 AGAAACGGTGTGGGTCTTGCCAGATACGACTTTTCAAGTTTAAAGACTTCTCTCTCC 964
Db 2234 AGAAACGGTGTGGGTCTTGCCAGATACGACTTTTCAAGTTTAAAGACTTCTCTCTCC 2293
QY 965 ACCACTGTATGAAGGACAGTGTCTCTCTCTGCGACCAACACAGGTGGATCTCCAA 1024
Db 2294 ACCACTGTATGAAGGACAGTGTCTCTCTCTGCGACCAACACAGGTGGATCTCCAA 2353
QY 1025 GGGATAGGAAGCCCTCGCGGAGAGTCTTTGTTTGGAGGCAAAAGCAGCAGCTGG 1084
Db 2354 GGGATAGGAGCCCTCGCGGAGAGTCTTTGTTTGGAGGCAAAAGCAGCAGCTGG 2413
QY 1085 AGGAGGACAGAAAGACACTGTATGTTTGGAAAGCTCGGAGGACGTGGAGGATGGC 1144
Db 2414 AGGAGGACAGAAAGACACTGTATGTTTGGAAAGCTCGGAGGACGTGGAGGATGGC 2473
QY 1145 AAGTGTCTGTGGGAGTTTCTGGGCAATGAATGCCAACAATGCTCTGTGCTGTGCGC 1204
Db 2474 AAGTGTCTGTGGGAGTTTCTGGGCAATGAATGCCAACAATGCTCTGTGCTGTGCGC 2533
QY 1205 GGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTGTGACCTCATCC 1264
Db 2534 GGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTGTGACCTCATCC 2593
QY 1265 TCATCCGGAATGCTCCAGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACGAGC 1324
Db 2594 TCATCCGGAATGCTCCAGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACGAGC 2653
QY 1325 AGGACAGTTGACTTCACTTTTGTGAATTTATCGGTTCAAGAAATTCAGTTTACGT 1384
Db 2654 AGGACAGTTGACTTCACTTTTGTGAATTTATCGGTTCAAGAAATTCAGTTTACGT 2713
QY 1385 CGAAGCACATGAGGATGAGGACGACCTCAAGAGGGGGGGAAGAGCGCTTTGGGC 1444
Db 2714 CGAAGCACATGAGGATGAGGACGACCTCAAGAGGGGGGGAAGAGCGCTTTGGGC 2773
QY 1445 ACATTTGAGGACGACCCCTCTGCTGTGACCGTCTCCAAAGCTCTCTGGAACCTGCG 1504
Db 2774 ACATTTGAGGACGACCCCTCTGCTGTGACCGTCTCCAAAGCTCTCTGGAACCTGCG 2833
QY 1505 ACGGGAGGTCTGTGACAGCCCTGCCATGAGGTGAGGTGAGTCCATGCCAGCTGTTGAC 1564
Db 2834 ACGGGAGGTCTGTGACAGCCCTGCCATGAGGTGAGGTGAGTCCATGCCAGCTGTTGAC 2893
QY 1565 TCTTTGACGAGGAATTTGAAGAAATTCGAGCCAGACTCACAAGCTGA 1614

Db 2894 TCTTTCGACGAGGAATTCGAGGATCCGAGCCAGACTCACAAGCTGA 2943

RESULT 12 AAS77728

ID AAS77728 standard; cDNA; 2241 BP.

XX AAS77728;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #13532.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR P-PSDB; ABG13541.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 1; SEQ ID NO 13532; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2241 BP; 453 A; 627 C; 700 G; 461 T; 0 U; 0 Other;

XX Query Match 57.7%; Score 931; DB 5; Length 2241;
XX Best Local Similarity 80.2%; Pred. No. 2.4e-231;
XX Matches 1273; Conservative 0; Mismatches 5; Indels 309; Gaps 3;

QY 140 CGGATGCTGCTGTGCTGCTATCTGAGATCATCGCGTTGAGGAAACACACGTTTCAG 199

Db 659 CAGATGCTGCTGTGCTGCTATCTGAGATCATCGCGTTGAGGAAACACACGTTTCAG 718

QY	200	GGAAACATCAAGGCAGTGGAAAAATGGCAGAAAAATGGAAAAACCTTACGCTTTTACAGTTC	259
DB	719	GGAAAAATCAAGGCAGTGGAAAAATGGCAGAAAAATGGAAAAACCTTACGCTTTTACAGTTC	778
QY	260	ACTGTGTAAAGAGAGACGACGCGCATCGCTGGAAAGTGGCGCAGAGTGACTTTCTGTGTGTC	319
DB	779	ACTGTGTAAAGAGAGACGACGCGCACCGCTGGAAAGTGGCGCAGAGTGACTTTCTGTGTGTC	838
QY	320	CAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCCCTGCGGGAGATGCTCGAGAACTGA	379
DB	839	CAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCCCTGCGGGAGATGCTCGAGAACTGA	898
QY	380	CGTCCAGACCAAGCAATTACTGTGATTTATCAACCCGTTTGGAGGAAAAAGCAAGGCA	439
DB	899	-----	898
QY	440	AGCGGATATAGAAAGAAAAAGTGGCACCATGTTCACCTTAGCCTCATCACCACCTGACA	499
DB	899	-----	898
QY	500	TCATCGTTACTGACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACA	559
DB	899	-----TTACTGAAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACA	952
QY	560	AATACGACGGCATCGTCTGTGTGGCGGAGATGGTATGTTTCAAGCGAGGTGCTGCACGGTC	619
DB	953	AATACGACGGCATCGTCTGTGTGGCGGAGATGGTATGTTTCAAGCGAGGTGCTGCACGGTC	1012
QY	620	TGATTGGAGGACGACAGAGAGCGCGGGTTCGACAGAAACCAACCCCGGGCTGTGCTGG	679
DB	1013	TGATTGGAGGACGACAGAGAGCGCGGGTTCGACAGAAACCAACCCCGGGCTGTGCTGG	1072
QY	680	TCCCCAGTAGCTCCGGAATTGGAATCATTTCCCGCA	714
DB	1073	TCCCCAGTAGCTCCGGAATTGGAATCATTTCCCGCAGAGCATGTGGGCCCGCGTGAGAAGC	1132
QY	715	-----GGGTCAAGGACATGCGTGTGTACTCCACCGTGGGCACACGAGCAGC	760
DB	1133	CTGTGTGCTTGGACCGGTCAACGACATGCGTGTGTACTCCACCGTGGGCACACGAGCAGC	1192
QY	761	CAGAAACCTCGGCGCTCATATCGTTT	789
DB	1193	CAGAAACCTCGGCGCTCATATCGTTTGTGCTGCTGCCCGAGGCTCGGAACCAACCCG	1252
QY	790	-----	789
DB	1253	CATCCCGCCATFACTGCTGTGGCAGTGGGCACGCGGACATGTGCTGGGCTGCCAGCTGT	1312
QY	790	-----GGGACAT	796
DB	1313	GGAAACGATGCTCTGTAGGCCTCGAGGCTTCAAGTCCAGAGTCAGAGCCCGGGGACT	1372
QY	797	CGCTGGCCATGAGTGTCTCTCAGTCCACACACAGCACACTCTTTCGCTACTCCGTGT	856
DB	1373	CGCTGGCCATGAGTGTCTCTCAGTCCACACACAGCACACTCTTTCGCTACTCCGTGT	1432
QY	857	CCCTGTGGGCTACGGTCTTCTACGGGACATCATCAAGGACAGTGAAGAAACCGTGGT	916
DB	1433	CCCTGTGGGCTACGGTCTTCTACGGGACATCATCAAGGACAGTGAAGAAACCGTGGT	1492
QY	917	TGGGTCTTGCAGATACGACTTTTCAAGTTTAAAGACTTCTCTCCACCACTGCTATG	976
DB	1493	TGGGTCTTGCAGATACGACTTTTCAAGTTTAAAGACTTCTCTCCACCACTGCTATG	1552
QY	977	AAGGACAGTGTCTCTCTCCCTGCACAAACACAGCTGGGATCTCCCAAGGATAGGAAGC	1036
DB	1553	AAGGACAGTGTCTCTCTCCCTGCACAAACACAGCTGGGATCTCCCAAGGATAGGAAGC	1612
QY	1037	CCTGCCGGGACGAGTGTCTTGTGTGAGGCAAGCAACGACAGCTCGAGGAGGAGCAGA	1096
DB	1613	CCTGCCGGGACGAGTGTCTTGTGTGAGGCAAGCAACGACAGCTCGAGGAGGAGCAGA	1672

Qy	1097	AGAAAGCACTGTATGTTTGGAAAGCTGCGGAGAGCACTGGAGGAGTGCGCAAGTCTGTCTGTG	1156
Db	1673	AGAAAGCACTGTATGTTTGGAAAGCTGCGGAGGACCTGGAGGAGTGGCAAGTCTGTCTGTG	1732
Qy	1157	GGAAAGTTTCTGGCCATCAATGCCACAAAATGTCCTGTGCTTGTGCGCGGAGCCCCCAGGG	1216
Db	1733	GGAAAGTTTCTGGCCATCAATGCCACAAAATGTCCTGTGCTTGTGCGCGGAGCCCCCAGGG	1792
Qy	1217	GCCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCTCATCCGGAAT	1276
Db	1793	GCCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCTCATCCGGAAT	1852
Qy	1277	GCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACGAGCAGGACCAAGTTTG	1336
Db	1853	GCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACGAGCAGGACCAAGTTTG	1912
Qy	1337	ACTTCACATTGTTGTAAGTTTATCGCGTCAAGAAATTCAGTTTACGTCGAAGCACATGG	1396
Db	1913	ACTTCACATTGTTGTAAGTTTATCGCGTCAAGAAATTCAGTTTACGTCGAAGCACATGG	1972
Qy	1397	AGGATGAGGACACGCGACCTCAAGGAGGGGGGAAGAGCGCTTTGGGCACATTTGCAGCA	1456
Db	1973	AGGATGAGGACACGCGACCTCAAGGAGGGGGGAAGAGCGCTTTGGGCACATTTGCAGCA	2032
Qy	1457	GCCACCCCTCTCTGCTGCACCGTCTCCAAACAGCTCTCTGGAACTCGACGGGAGGTCC	1516
Db	2033	GCCACCCCTCTCTGCTGCACCGTCTCCAAACAGCTCTCTGGAACTCGACGGGAGGTCC	2092
Qy	1517	TGCACAGCCCTGCCATCGAGGTCAGAG	1543
Db	2093	TGCACAGCCCTGCCATCGAGGTCAGAG	2119
RESULT 13			
ABL40822			
ID ABL40822 standard; cDNA; 979 BP.			
XX	ABL40822;		
AC			
XX			
DT	03-JUL-2002 (first entry)		
XX	Human sphingosine kinase-like protein encoding cDNA.		
XX	Human sphingosine kinase-like protein; intracellular signalling; gene;		
KW	cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;		
KW	autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	2..789	
FT		/*tag= a	
FT		/product= "sphingosine kinase-like protein"	
FT		/note= "start and stop codons are not indicated"	
XX			
XX	W0200228906-A2.		
PD			
XX	11-APR-2002.		
XX			
XX	05-OCT-2001; 2001WO-EP011516.		
XX			
PR	06-OCT-2000; 2000US-0238005P.		
PR	23-AUG-2001; 2001US-0314113P.		
XX			
XX	(FARB) BAYER AG.		
PA			
XX			
PI	Kossida S, Encinas J;		
XX			
DR	WPI; 2002-340094/37.		
DR	P-PSDE; ABB07854.		
XX			
PT	New reagent for modulating the activity of sphingosine kinase-like		
PT	protein polypeptide or polynucleotide and treating cancer, asthma,		

PT allergy, an autoimmune disease, or a central or peripheral nervous system disorder.

XX Claim 1; Fig 1; 120pp; English.

XX The invention relates to a human sphingosine kinase-like protein. The polypeptide can be expressed by standard recombinant methodology. The CC sphingosine kinase-like protein and gene can be used to regulate CC intracellular signalling and consequently cell proliferation and CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g. CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and CC peripheral nervous system disorders (e.g. Parkinson's disease). The CC present sequence represents the human sphingosine kinase-like protein CC encoding cDNA

XX Sequence 979 BP; 237 A; 259 C; 271 G; 212 T; 0 U; 0 Other;

Query Match 55.9%; Score 903; DB 6; Length 979;
Best Local Similarity 96.6%; Pred. No. 3.2e-224;
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;
QY 387 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAGCAAGCGGAT 446
Db 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAGCAAGCGGAT 60
QY 447 ATATCAAGAAAGTGGCACCACCTGTTACCTTAGCCTCCATCACCACCTGATCG- 505
Db 61 ATATCAAGAAAGTGGCACCACCTGTTACCTTAGCCTCCATCACCACCTGATCGG 120
QY 506 -----TTACTGAACATGCTTAATCAGGCCAAGGA 533
Db 121 TAACAAATTCATGTTAATCTGTAGAAAGTAATTAATGACATGCTAATCAGGCCAAGGA 180
QY 534 GACTCTGATGAGATTAAACATAGACAAATACAGCGGCATCGTCTGTGCGGGAGATGG 593
Db 181 GACTCTGATGAGATTAAACATAGACAAATACAGCGGCATCGTCTGTGCGGGAGATGG 240
QY 594 TATGTTACGAGGTTGCTGCAGGTTCTGATTTGGGAGAGCGCAGAGAGGCGCGGGTGA 653
Db 241 TATGTTACGAGGTTGCTGCAGGTTCTGATTTGGGAGAGCGCAGAGAGGCGCGGGTGA 300
QY 654 CCAGAACACACCCCGGCTGCTGCTCCAGTAGCTCCGGATTGGAATCATTTCCGCG 713
Db 301 CCAGAACACACCCCGGCTGCTGCTCCAGTAGCTCCGGATTGGAATCATTTCCGCG 360
QY 714 AGGGTCAACGGAAGTCTGTTTACTCCACCGTGGGACACGAGCGAGAAACCTCGGC 773
Db 361 AGGGTCAACGGAAGTCTGTTTACTCCACCGTGGGACACGAGCGAGAAACCTCGGC 420
QY 774 GCTGATATCGTTGTTGGGACTCGTGGCCATGATGTCTCAGTCCACCAACAG 833
Db 421 GCTGATATCGTTGTTGGGACTCGTGGCCATGATGTCTCAGTCCACCAACAG 480
QY 834 CACACTCTCTCGTACTCGTGTCTCTGCTGGGCTTCTACGGGACATCATCA 893
Db 481 CACACTCTCTCGTACTCGTGTCTCTGCTGGGCTTCTACGGGACATCATCA 540
QY 894 GGACAGTGAAGAAACGGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGAC 953
Db 541 GGACAGTGAAGAAACGGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGAC 600
QY 954 TTCTCTCTCCACCACTGTTATGAGGACAGTGTCTTCTCTCCCTGACCAACACCGGT 1013
Db 601 TTCTCTCTCCACCACTGTTATGAGGACAGTGTCTTCTCTCCCTGACCAACACCGGT 660
QY 1014 GGGATCTCAAGGATAGAACCTCGCGGACAGATGCTTTGTTTCAGGCAAGCAA 1073
Db 661 GGGATCTCAAGGATAGAACCTCGCGGACAGATGCTTTGTTTCAGGCAAGCAA 720
QY 1074 GCAGCAGCTGGAGGAGGACAGAAAGACACTGTATGTTTGGAGCTGCGGAGACGT 1133
Db 721 GCAGCAGCTGGAGGAGGACAGAAAGACACTGTATGTTTGGAGCTGCGGAGACGT 780

QY 1134 GGAGAGTGGCAAGTCGTCTCTGGGAAGTTTCTGSCCATCAATGCCAACAATGTCCTG 1193
Db 781 GGAGAGTGGCAAGTCGTCTCTGGGAAGTTTCTGSCCATCAATGCCAACAATGTCCTG 840
QY 1194 TGCTTGTGCGCGAGCCCCAGGGGCTCTCCCCGGCTGCCACTTTGGGAGACGGTCTTC 1253
Db 841 TGCTTGTGCGCGAGCCCCAGGGGCTCTCCCCGGCTGCCACTTTGGGAGACGGTCTTC 900
QY 1254 TGACCTCATCTCATCCGGAATGCTCAGGTTCAATTTCTGAGATTCTCATCAGGCA 1313
Db 901 TGACCTCATCTCATCCGGAATGCTCAGGTTCAATTTCTGAGATTCTCATCAGGCA 960
QY 1314 CACCAACAGCAGGACCAG 1332
Db 961 CACCAACAGCAGGACCAG 979
RESULT 14
AAS77730
ID AAS77730 standard; cDNA; 2186 BP.
XX
AC AAS77730;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13534.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
PD
XX 11-OCT-2001.
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG13543.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 13534; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2186 BP; 455 A; 637 C; 669 G; 425 T; 0 U; 0 Other;

Query Match 46.7%; Score 753.8; DB 5; Length 2186;
Best Local Similarity 99.7%; Pred. No. 2.9e-185;
Matches 755; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 790 GGGGACTCGTGGCCATGGATGTCTCTCAGTCCACACACAGCACACTCTTCGGTAC 849
Db 876 GGGGACTCGTGGCCATGGATGTCTCTCAGTCCACACACAGCACACTCTTCGGTAC 935

Qy 850 TCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAA 909
Db 936 TCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAA 995

Qy 910 CGGTGGTGGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACAC 969
Db 996 CGGTGGTGGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACAC 1055

Qy 970 TGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACAGGTGGATCTCCAAGGAT 1029
Db 1056 TGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACAGGTGGATCTCCAAGGAT 1115

Qy 1030 AGGAAGCCCTCCCGGAGGATGCTTTGTTGACGGAACAAAGCAGCAGCTGGAGAG 1089
Db 1116 AGGAAGCCCTCCCGGAGGATGCTTTGTTGACGGAACAAAGCAGCAGCTGGAGAG 1175

Qy 1090 GAGCAGAGAAGACACTGTATGTTGGTGGAGCTCGGAGGACGTGGAGGATGCAAGTC 1149
Db 1176 GAGCAGAGAAGACACTGTATGTTGGTGGAGCTCGGAGGACGTGGAGGATGCAAGTC 1235

Qy 1150 GTCTGTGGGAAGTTCTGGCCATCAATGCCACAAACATGCTGTGCTTGTGCGCGAGC 1209
Db 1236 GTCTGTGGGAAGTTCTGGCCATCAATGCCACAAACATGCTGTGCTTGTGCGCGAGC 1295

Qy 1210 CCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTCTGACCTCATCTCATC 1269
Db 1296 CCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTCTGACCTCATCTCATC 1355

Qy 1270 CGGAATGCTCCAGGTTCATTTTCTGAGATTTCTCATCAGGACACCAACACAGGAC 1329
Db 1356 CGGAATGCTCCAGGTTCATTTTCTGAGATTTCTCATCAGGACACCAACACAGGAC 1415

Qy 1330 CAGTTTGACTTCACTTTTGTGTAAGTTTATCGGTCAAGAAATTCAGTTTACGTGAG 1389
Db 1416 CAGTTTGACTTCACTTTTGTGTAAGTTTATCGGTCAAGAAATTCAGTTTACGTGAG 1475

Qy 1390 CACATGGAGGATGAGGACGACCTCAAGGAGGGGGAAGACGCTTTGGGCACATT 1449
Db 1476 CACATGGAGGATGAGGACGACCTCAAGGAGGGGGAAGACGCTTTGGGCACATT 1535

Qy 1450 TGCAGCAGCCACCCCTCTGCTGTGTCACCCGTCTCCACAGCTCTCTGSAATCGGACGG 1509
Db 1536 TGCAGCAGCCACCCCTCTGCTGTGTCACCCGTCTCCACAGCTCTCTGSAATCGGACGG 1595

Qy 1510 GAGTCTGTGACAGCCCTGCCATCGAGTCAAGTCC 1546
Db 1596 GAGTCTGTGACAGCCCTGCCATCGAGTCAAGAAC 1632

RESULT 15

AAS77731
ID AAS77731 standard; cDNA; 1570 BP.
XX
AC AAS77731;
XX
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13535.
XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-PSDB; ABG13544.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 13535; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1570 BP; 349 A; 446 C; 456 G; 319 T; 0 U; 0 Other;

Query Match 46.6%; Score 752.4; DB 5; Length 1570;
Best Local Similarity 99.9%; Pred. No. 5.8e-185;
Matches 753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 790 GGGGACTCGTGGCCATGGATGTCTCTCAGTCCACACACAGCACACTCTTCGGTAC 849
Db 695 GGGGACTCGTGGCCATGGATGTCTCTCAGTCCACACACAGCACACTCTTCGGTAC 754

Qy 850 TCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAA 909
Db 755 TCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAA 814

Qy 910 CGGTGGTGGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACAC 969
Db 815 CGGTGGTGGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACAC 874

Qy 970 TGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACAGGTGGATCTCCAAGGAT 1029
Db 875 TGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACAGGTGGATCTCCAAGGAT 934

Qy 1030 AGGAAGCCCTCCCGGAGGATGCTTTGTTGACGGAACAAAGCAGCAGCTGGAGAG 1089


```
|||||
Db 935 AGGAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCAAGCAGCAGCTGGAGGAG 994
QY 1090 GAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTGCGAGGACGTGGAGGAGTGGCAAGTC 1149
Db 995 GAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTGCGAGGACGTGGAGGAGTGGCAAGTC 1054
QY 1150 GTCTGTGGGAAGTTTCTGGCCCATCAATGSCCAAAACATGTCTGTGCTTGTGCGCCGGAGC 1209
Db 1055 GTCTGTGGGAAGTTTCTGGCCCATCAATGSCCAAAACATGTCTGTGCTTGTGCGCCGGAGC 1114
QY 1210 CCCAGGGGCTCTCCCGGCTGCCACATTGGGAGACGGGTCTTCTGACCTCATCTCATC 1269
Db 1115 CCCAGGGGCTCTCCCGGCTGCCACATTGGGAGACGGGTCTTCTGACCTCATCTCATC 1174
QY 1270 CGGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAACCAGCAGGAC 1329
Db 1175 CGGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAACCAGCAGGAC 1234
QY 1330 CAGTTTGACTTCACCTTTGTTGGAAGTTTATCGGTCAGAAATTCAGTTTACGTGGAAG 1389
Db 1235 CAGTTTGACTTCACCTTTGTTGGAAGTTTATCGGTCAGAAATTCAGTTTACGTGGAAG 1294
QY 1390 CACATGGAGGATGAGCAGCAGACCTCAAGGAGGGGGGGAAGAGCGCTTTGGGCACATT 1449
Db 1295 CACATGGAGGATGAGCAGCAGACCTCAAGGAGGGGGGGAAGAGCGCTTTGGGCACATT 1354
QY 1450 TGCAGCAGCCACCCCTCCTGTGCTGTGCACCCGTCTCCAACAGCTCCTGGAACTGCGACGGG 1509
Db 1355 TGCAGCAGCCACCCCTCCTGTGCTGTGCACCCGTCTCCAACAGCTCCTGGAACTGCGACGGG 1414
QY 1510 GAGGTCCTGACAGCCCTGCCATCGAGGTCAGAG 1543
Db 1415 GAGGTCCTGACAGCCCTGCCATCGAGGTCAGGG 1448
```

Search completed: September 5, 2005, 20:25:17
Job time : 939.933 secs

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	1612.4	99.9	4432	4	US-09-774-528-148	Sequence 148, App	
C	2	100.6	6.2	1084	4	US-09-270-767-15155	Sequence 15155, A
	3	99	6.1	2064	4	US-09-270-767-14306	Sequence 14306, A
C	4	69.8	4.3	498	4	US-09-893-737-59	Sequence 59, Appl
	5	67.8	4.2	1857	4	US-09-970-516-3	Sequence 3, Appl
6	67.8	4.2	2380	4	US-09-817-676A-13	Sequence 13, Appl	
7	65.6	4.1	901	4	US-09-270-767-30448	Sequence 30448, A	
8	61.2	3.8	2698	4	US-09-817-676A-11	Sequence 11, Appl	
9	51.2	3.2	1155	4	US-09-970-516-1	Sequence 1, Appl	
10	51.2	3.2	1205	4	US-09-959-897-1	Sequence 1, Appl	
11	51.2	3.2	1783	4	US-09-949-016-1155	Sequence 1155, Ap	
12	49.8	3.1	666	2	US-08-875-034A-1	Sequence 1, Appl	
13	49.6	3.1	1875	4	US-09-614-221A-399	Sequence 399, App	
14	49.2	3.0	1611	4	US-09-248-796A-1756	Sequence 1756, Ap	
15	48.4	3.0	1149	4	US-09-970-516-5	Sequence 5, Appl	
16	44.2	2.7	1533	1	US-09-205-258-90	Sequence 90, Appl	
C	17	43.8	2.7	7218	1	US-08-232-463-14	Sequence 14, Appl
	18	43.4	2.7	1050	4	US-09-252-991A-1269	Sequence 1269, Ap
19	43.4	2.7	1308	4	US-09-252-991A-1179	Sequence 1179, Ap	
C	20	43	2.7	3071	4	US-09-949-016-5812	Sequence 5812, Ap
C	21	43	2.7	62908	4	US-09-949-016-17554	Sequence 17554, A
	22	42.8	2.7	36519	3	US-08-923-137-2	Sequence 2, Appl
23	42.2	2.6	430	4	US-09-621-976-16656	Sequence 16656, A	
24	42.2	2.6	115963	4	US-09-949-016-12298	Sequence 12298, A	
25	41.2	2.6	1318	3	US-09-125-891-1	Sequence 1, Appl	
26	41.2	2.6	2636	3	US-09-125-891-3	Sequence 3, Appl	
27	40.8	2.5	1014	4	US-09-902-540-9500	Sequence 9500, Ap	

QY 181 GAGGAAACAGACGTTTCACGGGAAACATCAAGCGCAGTGGAAAAATGCGAGAAAAATGAAAAAG 240
DB 1384 GAGGAAACAGACGTTTCACGGGAAACATCAAGCGCAGTGGAAAAATGCGAGAAAAATGAAAAAG 1443
QY 241 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGACGACGCGCACCGCTGGAAAGTGGGCG 300
DB 1444 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGACGACGCGCACCGCTGGAAAGTGGGCG 1503
QY 301 CAGGTGACTTCTGTGTGTCAGAGAGAGACGCTGTGTCACTTGTGTGCTGCGACACCTTGGG 360
DB 1504 CAGGTGACTTCTGTGTGTCAGAGAGAGACGCTGTGTCACTTGTGTGCTGCGACACCTTGGG 1563
QY 361 GAGATGCTGGAGAGAGCTGACGCTCCAGACCAAGACATTTACTGTTATTTATCAACCCGTTT 420
DB 1564 GAGATGCTGGAGAGAGCTGACGCTCCAGACCAAGACATTTACTGGTATTTATCAACCCGTTT 1623
QY 421 GGAGGAAAGGACAAAGCAAGCGGATATATGAAAGAAAAGTGGCACCCTGTTTCACTTTA 480
DB 1624 GGAGGAAAGGACAAAGCAAGCGGATATATGAAAGAAAAGTGGCACCCTGTTTCACTTTA 1683
QY 481 GCCTCCATCACACTGACATCATGTGTACTGAAACATGCTATATCAGGCGCAAGGAGACTGTG 540
DB 1684 GCCTCCATCACACTGACATCATGTGTACTGAAACATGCTATATCAGGCGCAAGGAGACTGTG 1743
QY 541 TATGAGATTAAACATAGACAAATACAGCGGCATCGTCTGTGTGCGGGAGATGTTATGTTT 600
DB 1744 TATGAGATTAAACATAGACAAATACAGCGGCATCGTCTGTGTGCGGGAGATGTTATGTTT 1803
QY 601 AGCAGGTGCTGTCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGGTTCGACCAGAAC 660
DB 1804 AGCAGGTGCTGTCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGGTTCGACCAGAAC 1863
QY 661 CACCCCCGGCTGTGCTGGTCCCGAGTAGCCTTCGGATTGGAATCATTTCCCGCAGGGTCA 720
DB 1864 CACCCCCGGCTGTGCTGGTCCCGAGTAGCCTTCGGATTGGAATCATTTCCCGCAGGGTCA 1923
QY 721 ACGGACTCGCTGTGTTACTCCACCGTGGGACACGACGACGAGAACTCGCGCGCTGCAT 780
DB 1924 ACGGACTCGCTGTGTTACTCCACCGTGGGACACGACGACGAGAACTCGCGCGCTGCAT 1983
QY 781 ATCGTGTGTGGGACTCGCTGGCCATGATGTGCTCCAGTCCACCAACAGACACACTC 840
DB 1984 ATCGTGTGTGGGACTCGCTGGCCATGATGTGCTCCAGTCCACCAACAGACACACTC 2043
QY 841 CTTGCTACTCGGTGCTCCTGTGGCTACGGCTTCTACGGGACATCATCAAGGACAGT 900
DB 2044 CTTGCTACTCGGTGCTCCTGTGGCTACGGCTTCTACGGGACATCATCAAGGACAGT 2103
QY 901 GAGAGAAACGCTGTGGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTC 960
DB 2104 GAGAGAAACGCTGTGGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTC 2163
QY 961 TCCCAACCACTGCTATGAAGGACAGTGTCTTCTCCTCGCTGCAACACACGCTGGGATCT 1020
DB 2164 TCCCAACCACTGCTATGAAGGACAGTGTCTTCTCCTCGCTGCAACACACGCTGGGATCT 2223
QY 1021 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAAAGCAAGCAGGAG 1080
DB 2224 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAAAGCAAGCAGGAG 2283
QY 1081 CTGGAGGAGGACGAGAAAGACCTGTATGTTTGGAAAGCTGCGGAGGAGCTGGAGGAG 1140
DB 2284 CTGGAGGAGGACGAGAAAGACCTGTATGTTTGGAAAGCTGCGGAGGAGCTGGAGGAG 2343
QY 1141 TGGCAAGTGTGTGGGAAGTTTCTGGCCATCAATGCCAACAAACATGCTCTGTGCTTGT 1200
DB 2344 TGGCAAGTGTGTGGGAAGTTTCTGGCCATCAATGCCAACAAACATGCTCTGTGCTTGT 2403
QY 1201 CGCCGAGGCCAGGGGCTCTCCCGGCTGCCACTTTGGGAGAGCGGTCTTCTGACCTC 1260
DB 2404 CGCCGAGGCCAGGGGCTCTCTCCCGGCTGCCACTTTGGGAGAGCGGTCTTCTGACCTC 2463

QY 1261 ATCTCATTCGGGAAATGCTCCAGGTTCAATTTTCTGAGATTTTCTCATCAGGCACACCAAC 1320
DB 2464 ATCTCATTCGGGAAATGCTCCAGGTTCAATTTTCTGAGATTTTCTCATCAGGCACACCAAC 2523
QY 1321 CAGCAGACCAAGTTTGAATTTGACTTTTGTGTAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
DB 2524 CAGCAGACCAAGTTTGAATTTGACTTTTGTGTAAGTTTATCGCGTCAAGAAATTCAGTTT 2583
QY 1381 ACCTCGAAGCACATGGAGGATGAGCAGCAGCCTCAAGAGGGGGGAGAGAGCGCTTT 1440
DB 2584 ACCTCGAAGCACATGGAGGATGAGCAGCAGCCTCAAGAGGGGGGAGAGAGCGCTTT 2643
QY 1441 GGGCACATTTGACAGCACCCTCTGCTGTGTCACCGCTTCCAAACAGCTCTCTGGAAC 1500
DB 2644 GGGCACATTTGACAGCACCCTCTGCTGTGTCACCGCTTCCAAACAGCTCTCTGGAAC 2703
QY 1501 TGGACGGGAGGTCCTGCAAGCCCTGCCATCGAGGTCAAGGTCACACTGCGAGCTGTT 1560
DB 2704 TGGACGGGAGGTCCTGCAAGCCCTGCCATCGAGGTCAAGGTCACACTGCGAGCTGTT 2763
QY 1561 CGACTCTTTTCAGAGGAATTTGAGAGAAATCCGAAGCCAGACTCACACAGCTGA 1614
DB 2764 CGACTCTTTTCAGAGGAATTTGAGAGAAATCCGAAGCCAGACTCACACAGCTGA 2817
RESULT 2
US-09-270-767-15155/c
; Sequence 15155, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15155
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: *Drosophila melanogaster*
; US-09-270-767-15155
Query Match 6.2%; Score 100.6; DB 4; Length 1084;
Best Local Similarity 49.4%; Pred. No. 9.1e-17;
Matches 290; Conservative 0; Mismatches 294; Indels 3; Gaps 1;
QY 401 TGGTATTATCAACCCGTTTGGAGGAAAGGACAAAGCAAGCGGAGCTCAGACCTATGAGCGCCATG 969
DB 1028 TGGTCTTTATAAACCCCTATGGAGTCCAAAGGGGGAGCTCAGACCTATGAGCGCCATG 969
QY 461 TGGCACCACCTGTTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 520
DB 968 TGAGACCTATTTTTCAGCTTTGCGGGCGTGTAGCGCACCGTGTATCACCCTCAGAGGGCAA 909
QY 521 ATCAGGCGAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGCGATCCTCTGTG 580
DB 908 ACCAAGTGAAGGACATATCTTGTAGCCATGATCTGGAGGTATAGCTGCGGTTTGTCTGTG 849
QY 581 TCGCGGAGATGATGTTTTCAGCGAGGTGTGCAAGCTGTGATTTGGAGGACGAGGAGGAG 640
DB 848 TCGGAGGCGATGGCACCGTAGCAGAGGTATCAACCGGACTGATATTCCTGTAATGCGAG 789
QY 641 GCGCGGGGTTCAGACCAACCCCGGGCTGTGCTGCTCCAGTAGCTCCGATGCTCCGATG 700
DB 788 AGTTGGGACTGGAGCAACAGCGGCTCATAATATTCAGGACCGGCT---CTGCCAGTGG 732
QY 701 GAATCATTTCCCGCAGGGTCAACCGGACTGCGGTGTGTTACTCCACCGTGGGACACGAGCGACG 760
DB 731 GTGTGATTTCCGCTGGCAGGACCGACCACTTGGTATAGTATGACGCGGACGCGCGGATG 672
QY 761 CAGAAAACCTCGGCGCTGCATATCGTTTGTGGGAGACTCGTGGGCCATGGATGTGTCTTCAG 820

Db 671 TGAGGACAGCGGTATCCATGTGATTCGGCCAGCATCGGGATTGGATGTGTGACGTG 612
Qy 821 TCCACCAACAAGCAGCATCTTCGTACTCCGTGTCCCTGCTGGGTACGGCTTCTACG 880
Db 611 TGAGCAATGGCCAGTCCCTGCTCAGATTCTGTGCCAGTGTCTGAGCTACGGGTACCTGG 552
Qy 881 GGGACATCATCAAGGACAGTGAAGAAACGGTGGTGGTCTGCCAGATACGACTTTT 940
Db 551 GCGATGTGCCAGCCAGACGGAGAACTACCGCTGGATGGACCGCGCGGTACGAGTACA 492
Qy 941 CAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTG 987
Db 491 GTGGCTCAAGGCCCTTCTGTAATATCGGGCTATGAGCCGCAACTG 445

RESULT 3
US-09-270-767-14306
; Sequence 14306, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14306
; LENGTH: 2064
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14306

Query Match 6.1%; Score 99; DB 4; Length 2064;
Best Local Similarity 49.2%; Pred. No. 3.4e-16;
Matches 289; Conservative 0; Mismatches 295; Indels 3; Gaps 1;

Qy 401 TGGTATTATCAACCGTTTGGAGAAAGGCAAGCGGATATATGAAGAAAG 460
Db 620 TGGTCTTTATAAACCCCTATGGAGTGGCAAGCGGGAGCTCAGACCTATGAGCGCCATG 679
Qy 461 TGGCACCCTGTTACCTTTAGCTCCATCACCACCTGACATCATCTGTACTGATCAATGCTA 520
Db 680 TGAGACCTATTTTCAGCTTCCGGGTGAGCGGAGTACGCCAGCTGTATCACCCTAGAGGGCAA 739
Qy 521 ATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGGCATCTGTCTGTG 580
Db 740 ACCAAGTGAAGGACATCTTCTGAGCCATGATCTGGAGTATACGATGCGGTTTCTGTG 799
Qy 581 TCGGCGGAGATGGTATGTTTACGGAGGTCTGCAGGTCTGATTTGGGAGGACGACAGGA 640
Db 800 TCGGAGGCGATGGCAGGTAGCAGAGGTATCAACGGAGTGTATTTCCGTCAAATGCGAG 859
Qy 641 GCGCGGGGTCCACGACACACCCCGGGCTGTCTGTCTCCAGTAGCTCCGGATTG 700
Db 860 AGTGGGACTGGACGAAACAGCGGCCACCATATCTCAAGACCGGTCT--CTGCCAGTGG 916
Qy 701 GAATCATTTCCCGAGGGTCAAGGACTGCGTGTGTACTCCACCGTGGGCACGACGACG 760
Db 917 GTGTGATTTCCCGTGGCAGCAGCAGACCATTTGCGGTATAGTATGACGCGCAGCGGATG 976
Qy 761 CAGAAACCTTCGCGCTGCATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTCTCTCAG 820
Db 977 TGAGGACAGCGGTATCCATGTGATTTCTGGGCGAGCATCGGGGATTTGGATGTGTGAGTG 1036
Qy 821 TCCACCAACAAGCAGCATCTTCGTACTCCGTGTCCCTGCTGGGTACGGCTTCTACG 880
Db 1037 TGAGCAATGGCAGTCCCTGCTCAGATTCTGTGCCAGTGTCTGAGCTACGGGTACCTGG 1096
Qy 881 GGGACATCATCAAGGACAGTGAAGAAACGGTGGTGGTCTTGCAGATACGACTTTT 940
Db 1097 GCGATGTGCGACCCAGACGGAGAACTACCGCTGGATGGACCGCGCGGTACGAGTACA 1156

Qy 941 CAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAGGACAGTG 987
Db 1157 GTGGCGTCAAGGCGCTTCTGTAATATCGCGCTATGACGCCGAACTG 1203

RESULT 4
US-09-893-737-59/c
; Sequence 59, Application US/09893737
; Patent No. 6822082
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(498)
US-09-893-737-59

Query Match 4.3%; Score 69.8; DB 4; Length 498;
Best Local Similarity 91.4%; Pred. No. 1.1e-08;
Matches 74; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 713 CAGGTCACACGAGCTGCGTGTGTTACTCCACCGTGGCAGCAGCAGCAAACTCGG 772
Db 460 CAGGTCACACGAGCTGCGTGTGTTACTCCACCGTGGGACCGCGCGCAGAAACCGCGG 401

Qy 773 CGCTGCATATCGTTGTTGGGG 793
Db 400 CGCTGCATATCGTTTGTGGG 380

RESULT 5
US-09-970-516-3
; Sequence 3, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1857)
; OTHER INFORMATION:
US-09-970-516-3

Query Match 4.2%; Score 67.8; DB 4; Length 1857;
Best Local Similarity 45.4%; Pred. No. 7.7e-08;
Matches 292; Conservative 0; Mismatches 342; Indels 9; Gaps 1;

Qy 359 GGGAGATGCTGGAGAGCTGACGTCCAGACCAAGAGCATTTACTGTATTATCAACCGT 418
Db 401 GGGAGATCACCCCTGACCTGTACCTCGGCCCGCGGTGCTTCTATTGTGTCATCCCT 460

```
Qy 419 TTGAGGAGAAAGGACAGGCAAGCGGATATATGAAAGAAAGTGGCCACCAGTGTTCACCT 478
Db 461 TTGGGGGTCGGGGCTGGCCCTGGCAGTGGTGTAAAGAACACAGTCTTCCCATGATCTCTG 520
Qy 479 TAGCCTCCATCACCACCTGACATCATCGTTACTGTGAACATGCTAAATCAGGCCAAGGAGACTC 538
Db 521 AAGCTGGGCTGTCTCTTCAACCTCATCCAGACAGAACGACAGAACCCGCGGAGCTGG 580
Qy 539 TGTATGAGATTAAACATAGACAAATACGACGGCATCTGTCTGTGGGGAGATGGTATGT 598
Db 581 TCCAGGGGCTGAGCCTGAGTGAGTGGGATGGCATCTGTCAAGTCTCGGGAGACGGGCTGC 640
Qy 599 TCAGCGAGGTCTGTCAGCGTCTGATTTGGGAGGACGACAGAGAGCGCCGGGGTGCACAGA 658
Db 641 TCCATGAGGTCTGTAACGGGCTCTAGATCGCCCTGACTGGGAGGAAAGCTGTGAAGATG- 699
Qy 659 ACCACCCCGGGCTGTGTGTCTCCAGTAGACCTCCGGATTGGAAATCATTTCCCGCAGGCT 718
Db 700 -----CCTGTGGGCATCTCTCCCTCGGCTCGGGCAACGCGCTGGCCGGAGCAGTGA 751
Qy 719 CAACGGACTGGTGTGTACTCCACCGTGGGCACACGACGACGAGAAACCTTCGGCGCTGC 778
Db 752 ACCAGACGGGGATTTGAGCCAGCCCTGGGCTCGACCTTGTGCTCAACTGTCTCACTGT 811
Qy 779 ATATCGTTTGGGACTCGTGGCCATGGATGTCTCTCAGTCCACACACACAGCACAC 838
Db 812 TGCTGTGCGGGGTGGTGGCCACCACTGGACCTGCTCTCGTGACGCTGGCCCTCGGGCT 871
Qy 839 TCCTTCGCTACTCGTGTCTCTGTGGCTAGGCTTCTAGCGGAGACATCATCAAGACA 898
Db 872 CCCGCTGTTCTCTCTCTGTGTGGCTTGGGCTTGGGCTTCTCAGATGTGGATATCCAGA 931
Qy 899 GTGAGAGAAACGGTGTGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCC 958
Db 932 CGAGCGCTTCAGGGCTTGGGCTTGGGCTTGGGCTTCTCAGTCCAGCTGGGCAACGCTG 991
Qy 959 TCTCCCACTGCTATGAAGGACAGTGTCTTCTTCTCCCTGC 1001
Db 992 CCACACTGCACACTACCGGAGCGCTCTCTACCTCCCGC 1034
```

RESULT 6

US-09-817-676A-13

; Sequence 13, Application US/09817676A

; Patent No. 6800470

; GENERAL INFORMATION:

; APPLICANT: Spiegel, Sarah

; APPLICANT: Kohama, Takafumi

; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,

; FILE REFERENCE: Expression and Methods of Use Thereof

; FILE REFERENCE: 00170/HG

; CURRENT APPLICATION NUMBER: US/09/817,676A

; CURRENT FILING DATE: 2001-03-26

; PRIOR APPLICATION NUMBER: US 60/194,318

; PRIOR FILING DATE: 2000-04-03

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13

; LENGTH: 2380

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (7)..(1860)

; PUBLICATION INFORMATION:

; TITLE: Molecular cloning and functional characterization of a

; TITLE: novel mammalian sphingosine kinase type 2 isoform

; JOURNAL: J. Biol. Chem.

; VOLUME: 275

; ISSUE: 26

; PAGES: 19513-19520

; DATE: 2000-06-30

; DATABASE ACCESSION NUMBER: AF245447

; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-13

Query Match 4.2%; Score 67.8; DB 4; Length 2380;

Best Local Similarity 45.4%; Pred. No. 8.8e-08;

Matches 292; Conservative 0; Mismatches 342; Indels 9; Gaps 1;

```
Qy 359 GGGAGATGCTGGAGAAAGCTGACGTCCAGACCAAAAGCAATTTACTGGTATTTATCAACCGT 418
Db 407 GGGAGATCACCCCTGACCTGCTACCTCGGCGGCCCGTTCCTTATTGGTCAATCCCT 466
Qy 419 TTGGAGGAAAAAGACAAGGCAAGGGATATATGAAGAAAAGTGGACCACTGTTCACT 478
Db 467 TTGGGGTTCGGGGCTGGCCTGGCAGTGGTGTGAAGAACCAACGCTGTTCCATGATCTGT 526
Qy 479 TAGCTCCATCACCACTGACATCATCGTTACTGAAATGCTAATCAGGCCCAAGGAGACTC 538
Db 527 AAGCTGGGCTGTCTTCAACCTCATCCAGACAGAACGACAGAACCCGCGGAGCTGG 586
Qy 539 TGTATGAGATTAAACATAGACAAATACGACGGCATCTGTCTGTGGCGGAGATGGTATGT 598
Db 587 TCCAGGGGCTGAGCCTGAGTGAGTGGATGGCATCGTCAAGTCTCGGAGACGGGCTGC 646
Qy 599 TCAGCGAGGTGCTGCACGGTCTGATTTGGGAGGACGACAGAGAGCGCCGGGGTGCACAGA 658
Db 647 TCCATGAGGTGCTGAACGGGCTCTTAGATCGCCCTGACTGGGAGGAAGCTGTGAAGATG- 705
Qy 659 ACCACCCCGGGCTGTGTCTTCCCAAGTAGCCTCCGGATTGGAAATCATTTCCCGCAGGCT 718
Db 706 -----CCTGTGGGCATCTCTCCCTCGGCTCGGCAACGCGCTGGCCGGAGCAGTGA 757
Qy 719 CAACGGACTCGTGTGTACTCCACCGTGGGCACACGACGACGAGAAACCTTCGGCGCTGC 778
Db 758 ACCAGACGCGGGATTTGAGCCAGCCCTGGGCTCGACCTTGTGCTCAACTGTCTCACTGT 817
Qy 779 ATATCGTTTGGGACTCGTGGCCATGGATGTCTCTCAGTCCACACACACAGCACAC 838
Db 818 TGCTGTGCGGGGTGGTGGCCACCCACTGACCTGCTCTCCGTGACGCTGGCCCTCGGGCT 877
Qy 839 TCCCTTCGCTACTCGTGTCTCTGTGGCTACGGCTTCTACGGGGACATCATCAAGGACA 898
Db 878 CCCGCTGTTCTCTCTCTGTGTGGCCTTGGGCTTCTGATGTGGATATCCAGA 937
Qy 899 GTGAGAGAAACGGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACCTTCC 958
Db 938 GCGAGCGCTTCAGGGCCTTGGGCAAGTCCCGCTTCACACTGGGCAACGCTGCTGGGCTCG 997
Qy 959 TCTCCCACTGCTATGAAGGACAGTGTCTTCTTCTCCCTGC 1001
Db 998 CCACACTGCACACTACCGGAGCGCTCTCTACCTCCCGC 1040
```

RESULT 7

US-09-270-767-30448

; Sequence 30448, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 30448

; LENGTH: 901

; TYPE: DNA

; ORGANISM: *Drosophila melanogaster*

US-09-270-767-30448

Query Match

Best Local Similarity 50.0%; Score 65.6; DB 4; Length 901;

Matches 223; Conservative 0; Mismatches 214; Indels 9; Gaps 2;

Query Match 3.8%; Score 61.2; DB 4; Length 2698;
Best Local Similarity 53.3%; Pred. No. 5.5e-06;
Matches 129; Conservative 0; Mismatches 113; Indels 0

; Patent No. 6730480

```
; GENERAL INFORMATION:
; APPLICANT: PITSON, Stuart M
; APPLICANT: Brian, WATTENBERG W
; APPLICANT: Pu, XIA
; APPLICANT: Richard, D'ANDREA J
; APPLICANT: Jennifer, BAMBLE R
; APPLICANT: Mathew, VADAS A
; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
; FILE REFERENCE: PITSON=1
; CURRENT APPLICATION NUMBER: US/09/959,897
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00457
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: AU PQ 0339
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: AU PQ 1504
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1205
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (33)..(1184)
; OTHER INFORMATION:
US-09-959-897-1

Query Match      3.2%; Score 51.2; DB 4; Length 1205;
Best Local Similarity 49.6%; Pred. No. 0.0018;
Matches 131; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

Qy 398 TACTGGTATTATCAACCCGTTTGGAGGAAAGGACAGGCGGATATATGAAGAA 457
Db 82 TGCTGGTGCTGCTGAACCCGCGCGGCGGAGGCGCTTGAGCTCTTCGGAGTC 141
Qy 458 AAGTGGCACCACCTGTTTCACTTACCTTAGCCTCCATCACCACCTGACATCATCGTTTACTGAACATG 517
Db 142 ACGTGACGCCCTTTTGGCTGAGGCTGAATCTCCTTCACGCTGATGCTCACTGAGCGGC 201
Qy 518 CTAATCAGGCGCAAGAGACTCTGTATGAGATTACATAGACAATAACAGCGGATCGTCT 577
Db 202 GGAACACGCGGGAGCTGGTGGCGTGGAGGAGCTGGGCGGCTGGGAGCGCTCTGGTGG 261
Qy 578 GTGTCGGCGGAGATGGTATGTTTACGCGAGGTGCTGACGCTGCTGATTGGGAGGACGAGA 637
Db 262 TCATGCTTGGAGACGGGCTGATGACGAGGTGGTGAACGGGCTCATGAGCGGCTGACT 321
Qy 638 GGAGCGCGGGGTCGACCAGAAC 661
Db 322 GGGAGACCGCCATCCAGAAGCCCC 345

RESULT 11
US-09-949-016-1155
; Sequence 1155, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 1155
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1155

Query Match      3.2%; Score 51.2; DB 4; Length 1783;
Best Local Similarity 49.6%; Pred. No. 0.0022;
Matches 131; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

Qy 398 TACTGGTATTATCAACCCGTTTGGAGGAAAGGACAGGCGGATATATGAAGAA 457
Db 386 TGCTGGTGCTGCTGAACCCGCGCGGCGGAGGCGCTTGAGCTCTTCGGAGTC 445
Qy 458 AAGTGGCACCACCTGTTTCACTTACCTTAGCCTCCATCACCACCTGACATCATCGTTTACTGAACATG 517
Db 446 ACGTGACGCCCTTTTGGCTGAGGCTGAATCTCCTTCACGCTGATGCTCACTGAGCGGC 505
Qy 518 CTAATCAGGCGCAAGAGACTCTGTATGAGATTACATAGACAATAACAGCGGATCGTCT 577
Db 506 GGAACACGCGGGAGCTGGTGGCGTGGAGGAGCTGGGCGGCTGGGAGCGCTCTGGTGG 565
Qy 578 GTGTCGGCGGAGATGGTATGTTTACGCGAGGTGCTGACGCTGCTGATTGGGAGGACGAGA 637
Db 566 TCATGCTTGGAGACGGGCTGATGACGAGGTGGTGAACGGGCTCATGAGCGGCTGACT 625
Qy 638 GGAGCGCGGGGTCGACCAGAAC 661
Db 626 GGGAGACCGCCATCCAGAAGCCCC 649

RESULT 12
US-08-875-034A-1
; Sequence 1, Application US/08875034A
; Patent No. 5968796
; GENERAL INFORMATION:
; APPLICANT: Bieseler, Barbara; Reinemer, Peter; Hain,
; APPLICANT: Rudiger; Mann, Karlheinz; Reif, Hans-Jorg; and
; APPLICANT: Thomzik, Jurgen Ernst
; TITLE OF INVENTION: DIOXYRIBONUCLEIC ACID CODING FOR
; GLUTATHIONE-S-TRANSFERASE AND ITS
; TITLE OF INVENTION: USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sprung Kramer Schaefer & Briscoe
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.5
; SOFTWARE: WordPerfect 3.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,034A
; FILING DATE: 16-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/00068
; FILING DATE: 10-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 01 840.0
; FILING DATE: 23-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: Bayer 9895-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; INFORMATION FOR SEQ ID NO: 1;
```


SEQUENCE CHARACTERISTICS:
LENGTH: 666 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..663
US-08-875-034A-1

Query Match 3.1%; Score 49.8; DB 2; Length 666;
Best Local Similarity 53.3%; Pred. No. 0.0031;
Matches 105; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 7 GCGACGGGGCGGCGGAGCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGCTGCGCC 66
DB 265 GCGACGGGTGCGGCGGAGCTGGAGGTGGCTGGAGGTGGAGTGGACCACTTCCAC 324
QY 67 GTGACCTGGAGCCCGCGGCGCTCTGCTGCGCTGTGCGGAGCCCGGGCGCGGAGCC 126
DB 325 CGGAACGGCTGCGCGCTGTGCTTCCAGCTGTCTGCTGAGCGCTCTCTGGGCGGCGCC 384
QY 127 GCGGCCCCGCGCGGATGCTGCTGTGCTGTATCTGAGATCATCGCGTTGAGAA 186
DB 385 GACGCGCGGCTGTGGAGAACACGCGGAGCAGCTCGCCAGGTGCTCGACGTGTACGAG 444
QY 187 ACAGACGTTTCACGGGA 203
DB 445 GCGCACCTGGCCCGCAA 461

RESULT 13
US-09-614-221A-399
Sequence 399, Application US/09614221A
Patent No. 6723837
GENERAL INFORMATION:
APPLICANT: Karunanandaa, Balasulojini
APPLICANT: Yu, Jaehyuk
APPLICANT: Kishore, Ganesh M.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM
FILE REFERENCE: 16516.075
CURRENT APPLICATION NUMBER: US/09/614,221A
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/142,981
PRIOR FILING DATE: 1999-07-12
NUMBER OF SEQ ID NOS: 626
SEQ ID NO 399
LENGTH: 1875
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-399
Query Match 3.1%; Score 49.6; DB 4; Length 1875;
Best Local Similarity 49.6%; Pred. No. 0.006;
Matches 127; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 398 TACTGTATTATCAACCGTTTGGAGGAAAGGACAGGCGGATATGAAAGAA 457
DB 686 TATTAGTCATTATTAATCCCAACGGTGGTAAAGGTACTGCTAAAAATTTATTCCTGACAA 745
QY 458 AAGTGGCCACCTGTTTCACTTACCTTCATCCACCTGACATCATCTGTTACTGAACATG 517
DB 746 AAGCAAGGCCATACTAGTGGAAAGTGGTGCAAAATAGAAATTCATACACAAAATATG 805
QY 518 CTAATCAGCCCAAGGAGACTCTGTATGAGATTAACATAGACAAATACGACGGCATCGTCT 577
DB 806 CCGGTACGCCCATCATATTTGCCAAGATTAGATATCAGCAAAATACGATACCATTTGCAT 865
QY 578 GTGTCCGGCGGAGATGGTATGTTCCAGGAGGTGCTCCACGGTCTGATTGGGAGGACGAGA 637

DB 866 GTGCTCGGTGATGTTATTCATACGAGTAATTAATGCGCTTTATAGAACCCGACA 925
QY 638 GGAGCGCGGGGTCCA 653
DB 926 GAGTGGATGCGTTCAA 941
RESULT 14
US-09-248-796A-1756
Sequence 1756, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 1756
LENGTH: 1611
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-1756
Query Match 3.0%; Score 49.2; DB 4; Length 1611;
Best Local Similarity 51.4%; Pred. No. 0.007;
Matches 114; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 396 TTTACTGTATTATCAACCGTTTGGAGGAAAGGACAAAGCGGATATATGAAAG 455
DB 492 TATTTGGTATTGATTAACCGCATGCGGCCAAGGACACGCCAAACTATTTACAAAA 551
QY 456 AAAAGTGGCACCACTGTTACCTTAGCTCCATCCACTGACATCATCGTTACTGAAACA 515
DB 552 TAAAACTTACCAATATTACAAGCGGCTGCTGCTAAATGTTACGTATTTTGAAGAACTAAAATA 611
QY 516 TGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAACATAGACAAATACGACGGCATCGT 575
DB 612 TCATGGACACGCCACTGAGATTGCGCGTAGCTAGATGTCAATGATTATGATATAATTTGT 671
QY 576 CTGTGTCGGCGAGATGTTATGTTTTCAGCGAGTCTGCACGG 617
DB 672 TTGTTGTTCTGGCGATGGGATACCTCATGAAGTTTCAATGG 713
RESULT 15
US-09-970-516-5
Sequence 5, Application US/09970516
Patent No. 6610534
GENERAL INFORMATION:
APPLICANT: No. 6610534artis AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
TITLE OF INVENTION: polynucleotides encoding sphingosine kinases
FILE REFERENCE: 4-31617
CURRENT APPLICATION NUMBER: US/09/970,516
CURRENT FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1149
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1149)
OTHER INFORMATION:
US-09-970-516-5

	Query Match	100.0%	Score 1614;	DB 10;	Length 1614;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 1614;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0
Qy	1	ATGGGGGCGACGGGGGGCGGGAGCCGCTGCATTCCTGTGTGGGTGAACGACGAGCGC	60		
Db	1	ATGGGGGCGACGGGGGGCGGGAGCCGCTGCATTCCTGTGTGGGTGAACGACGAGCGC	60		


```
QY 181 GAGGAAACAGAGCTTACCGGAAACATCAAGGAGTGGAAATGCGAGAAATGGAAG 240
DB 181 GAGGAAACAGAGCTTACCGGAAACATCAAGGAGTGGAAATGCGAGAAATGGAAG 240
QY 241 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGACGACGCGTGGAAAGTGGCG 300
DB 241 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGACGACGCGTGGAAAGTGGCG 300
QY 301 CAGGTGACTTTCTGTGTCCAGAGGAGCAGTGTGTCACTTGTGTGCGAGACCTGCGG 360
DB 301 CAGGTGACTTTCTGTGTCCAGAGGAGCAGTGTGTCACTTGTGTGCGAGACCTGCGG 360
QY 361 GAGATGCTGGAGAGCTGAGCTCCAGACCAAGAGATTTACTGGTATTTATCAACCGTTT 420
DB 361 GAGATGCTGGAGAGCTGAGCTCCAGACCAAGAGATTTACTGGTATTTATCAACCGTTT 420
QY 421 GGAGGAAAGAGCAAGGCAAGCGGATATATGAAGAAAGTGGCACCACTGTTTCACTTGA 480
DB 421 GGAGGAAAGAGCAAGGCAAGCGGATATATGAAGAAAGTGGCACCACTGTTTCACTTGA 480
QY 481 GCCTCCATCACCCTGACATCATCGTTACTGAAACATGCTTAATCAGGCCAAGGAGACTCTG 540
DB 481 GCCTCCATCACCCTGACATCATCGTTACTGAAACATGCTTAATCAGGCCAAGGAGACTCTG 540
QY 541 TATGAGATTAAATAGACAAATACGACGCGATCGTCTGTCTCGGCGGAGATGTTATGTT 600
DB 541 TATGAGATTAAATAGACAAATACGACGCGATCGTCTGTCTCGGCGGAGATGTTATGTT 600
QY 601 AGCGAGGTGCTCAGCGTCTGATTTGGAGGACGACGAGAGCGCGGGTCCAGCAGAAC 660
DB 601 AGCGAGGTGCTCAGCGTCTGATTTGGAGGACGACGAGAGCGCGGGTCCAGCAGAAC 660
QY 661 CACCCCCGGGCTGTGTGTGTCCTCCAGTACCTCCGATTTGGAATCAITCCCGCAGGGTCA 720
DB 661 CACCCCCGGGCTGTGTGTGTCCTCCAGTACCTCCGATTTGGAATCAITCCCGCAGGGTCA 720
QY 721 ACAGGATCGGTGTGTACTTCAACCGTGGGACGACGAGAGCGAGAACTCTGGCGGTGAT 780
DB 721 ACAGGATCGGTGTGTACTTCAACCGTGGGACGACGAGAGCGAGAACTCTGGCGGTGAT 780
QY 781 ATCGTGTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAGCACATC 840
DB 781 ATCGTGTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAGCACATC 840
QY 841 CTTGCTACTCCTGCTGCTGGGCTACCGCTTCTACGGGACATCATCAAGGACAGT 900
DB 841 CTTGCTACTCCTGCTGCTGGGCTACCGCTTCTACGGGACATCATCAAGGACAGT 900
QY 901 GAGAAAGACGCTGTGGTCTTGCAGATACGACTTTTCAAGTTTAAAGACCTTCCTC 960
DB 901 GAGAAAGACGCTGTGGTCTTGCAGATACGACTTTTCAAGTTTAAAGACCTTCCTC 960
QY 961 TCCCACTACTGTATGAAGGACAGTGTCTTCTCCCTGCAACACACAGCTGGGATCT 1020
DB 961 TCCCACTACTGTATGAAGGACAGTGTCTTCTCCCTGCAACACACAGCTGGGATCT 1020
QY 1021 CCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGAGGCAAGCAAGCAGCAG 1080
DB 1021 CCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGAGGCAAGCAAGCAGCAG 1080
QY 1081 CTGGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGTGAGGAG 1140
DB 1081 CTGGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGTGAGGAG 1140
QY 1141 TGGCAAGTCTGTGTGGAGTTTCTGGCCATCAATGCCAACAATGCTGTGCTTGT 1200
DB 1141 TGGCAAGTCTGTGTGGAGTTTCTGGCCATCAATGCCAACAATGCTGTGCTTGT 1200
QY 1201 CGCCGAGGCCAGGAGGCTCTCCCGGCTGCCACTTTGGGAGAGCGGTCTTCTGACCTC 1260
DB 1201 CGCCGAGGCCAGGAGGCTCTCCCGGCTGCCACTTTGGGAGAGCGGTCTTCTGACCTC 1260
```

```
QY 1261 ATCTCTATCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAAC 1320
DB 1261 ATCTCTATCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAAC 1320
QY 1321 CAGCAGGACCAAGTTGACTTTCACCTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
DB 1321 CAGCAGGACCAAGTTGACTTTCACCTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
QY 1381 ACCTGGAAGCACTGGAGGATGAGGACAGCGACTTCAAGGAGGGGGGAAAGCGCTTT 1440
DB 1381 ACCTGGAAGCACTGGAGGATGAGGACAGCGACTTCAAGGAGGGGGGAAAGCGCTTT 1440
QY 1441 GGGCACATTTCCAGCAGCCACCTCTGCTGCTGCACCGTCTCAAACAGCTCTCTGGAAC 1500
DB 1441 GGGCACATTTCCAGCAGCCACCTCTGCTGCTGCACCGTCTCAAACAGCTCTCTGGAAC 1500
QY 1501 TGGACGGGGAGGCTCTGCACAGCCCTGCCATCGAGGTGAGAGTCCACTGCGAGCTGTT 1560
DB 1501 TGGACGGGGAGGCTCTGCACAGCCCTGCCATCGAGGTGAGAGTCCACTGCGAGCTGTT 1560
QY 1561 CGACTCTTTGACGAGGAATTTGAAGAGAAATCCGAAGCCAGACTCACAGCTGA 1614
DB 1561 CGACTCTTTGACGAGGAATTTGAAGAGAAATCCGAAGCCAGACTCACAGCTGA 1614

RESULT 3
US-09-969-896-16
; Sequence 16, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: Kinase-Like Protein
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIORITY FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 4413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-16

Query Match 100.0%; Score 1614; DB 10; Length 4413;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGGCGACGGGGGGCGGAGCCGCTGCAATCCGTGCTGGGTCAAGCAGCAGCGC 60
DB 76 ATGGGGGGCGACGGGGGGCGGAGCCGCTGCAATCCGTGCTGGGTCAAGCAGCAGCGC 135
QY 61 TGGCGCTGAGCCCTGGAGCCCGCGCGGCTCTGCTGCGCTGCTGGCGAGCCCGGGGCC 120
DB 136 TGGCGCTGAGCCCTGGAGCCCGCGCGGCTCTGCTGCGCTGCTGGCGAGCCCGGGGCC 195
QY 121 GGAGCGGCGCCCGCGGCGGATGCTGCTGCTGCTGCTGATCTGAGATCATCGCCGTT 180
DB 196 GGAGCGGCGCCCGCGGCGGATGCTGCTGCTGCTGCTGATCTGAGATCATCGCCGTT 255
QY 181 GAGGAAACAGAGCTTACGGGAAACATCAAGGAGTGGAAATGCGACAAATGGAAG 240
DB 256 GAGGAAACAGAGCTTACGGGAAACATCAAGGAGTGGAAATGCGACAAATGGAAG 315
QY 241 CTTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGCACCGCTGGAAGTGGCG 300
DB 316 CTTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGCACCGCTGGAAGTGGCG 375
QY 301 CAGGTGACTTTCTGTGGTGTCCAGAGAGCAGCTGTGTCTACTTGTGCTGACAGACCTTCGCG 360
```

```
Db 376 CAGGTGACTTTCTGGTGTCCAGAGAGCAGCTGTGTCACTTGTGGCTGCAGACCCCTCGG 435
Qy 361 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTTACTGGTATTATCAACCCGTTT 420
Db 436 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTTACTGGTATTATCAACCCGTTT 495
Qy 421 GGAGGAAAGACAAAGCAAGCGGATATATGAAGAAAGTGGCAACCACTGTTTCACTTTA 480
Db 496 GGAGGAAAGACAAAGCAAGCGGATATATGAAGAAAGTGGCAACCACTGTTTCACTTTA 555
Qy 481 GCCTCCATCACCACCTGACATCATCGTTACTGAAACATGCTAATCAGGCAAGGAGACTGTG 540
Db 556 GCCTCCATCACCACCTGACATCATCGTTACTGAAACATGCTAATCAGGCAAGGAGACTGTG 615
Qy 541 TATGAGATTACATAGACAAATACAGCGGCATGCTGTGTGCGGGAGATGGTATGTTT 600
Db 616 TATGAGATTACATAGACAAATACAGCGGCATGCTGTGTGCGGGAGATGGTATGTTT 675
Qy 601 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGAGAGGAGCGCGGGTGCACCAAGAC 660
Db 676 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGAGAGGAGCGCGGGTGCACCAAGAC 735
Qy 661 CACCCCGGGCTGTCTGGTCCCAAGTAGCCTCCGGATTGGAATCATTTCCCGCAGGGTCA 720
Db 736 CACCCCGGGCTGTCTGGTCCCAAGTAGCCTCCGGATTGGAATCATTTCCCGCAGGGTCA 795
Qy 721 ACAGACTCGTGTGTTACTTCCACGCTGGGACACGAGCGAGCAACCTCGGCGCTGCAT 780
Db 796 ACAGACTCGTGTGTTACTTCCACGCTGGGACACGAGCGAGCAACCTCGGCGCTGCAT 855
Qy 781 ATCGTTGTTGGGACTCGCTGGCATGATGTCTCAGTCCACCAACAGACACACTC 840
Db 856 ATCGTTGTTGGGACTCGCTGGCATGATGTCTCAGTCCACCAACAGACACACTC 915
Qy 841 CTTGCTACTCCGTGCTCGTGGCTACCGGCTTCTACGGGGACATCATCAAGGACAGT 900
Db 916 CTTGCTACTCCGTGCTCGTGGCTACCGGCTTCTACGGGGACATCATCAAGGACAGT 975
Qy 901 GAGAGAACCGTGGTGGGTCTGCCAGATACGACTTTTTCAGTTTAAAGACCTTCTC 960
Db 976 GAGAGAACCGTGGTGGGTCTGCCAGATACGACTTTTTCAGTTTAAAGACCTTCTC 1035
Qy 961 TCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCGCAACACACCGTGGGATCT 1020
Db 1036 TCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCGCAACACACCGTGGGATCT 1095
Qy 1021 CCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGAG 1080
Db 1096 CCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGAG 1155
Qy 1081 CTGGAGGAGGACAGAAAGCACTGTATGGTTTGGAGCTGCGGAGGACGTGGAGGAG 1140
Db 1156 CTGGAGGAGGACAGAAAGCACTGTATGGTTTGGAGCTGCGGAGGACGTGGAGGAG 1215
Qy 1141 TGGCAAGTCTGTGTGGAAAGTTTCTGCCATCAATGCAACAAATCTCTGTGCTTGT 1200
Db 1216 TGGCAAGTCTGTGTGGAAAGTTTCTGCCATCAATGCAACAAATCTCTGTGCTTGT 1275
Qy 1201 CGCGGAGCCCCAGGGGCTCTCCCCGCTGCCACTTTGGGAGACGGGTCTTCTGACCTC 1260
Db 1276 CGCGGAGCCCCAGGGGCTCTCCCCGCTGCCACTTTGGGAGACGGGTCTTCTGACCTC 1335
Qy 1261 ATCCTCATCGGAAATGCTCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAAC 1320
Db 1336 ATCCTCATCGGAAATGCTCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAAC 1395
Qy 1321 CAGCAGGACCACTTTGACTTCACTTTTGTGAAGTTTATCGGTCAGAAATTCAGTTT 1380
Db 1396 CAGCAGGACCACTTTGACTTCACTTTTGTGAAGTTTATCGGTCAGAAATTCAGTTT 1455
Qy 1381 ACGTCCAAAGCACATGGAGGATGAGACAGCGACTCAAGGAGGGGGGAAAGCGCTTT 1440
```

```
Db 1456 ACCTCGAAGCACATGGAGGATGAGGACAGCGA CTTCAAGAGGGGGGGAAGACGGCTTT 1515
Qy 1441 GGGCACATTTGCAGCAGCCACCCTCTCTGTGTGTCACCCGTCTTCAACAGACTCTCTGAAC 1500
Db 1516 GGGCACATTTGCAGCAGCCACCCTCTCTGTGTGTCACCCGTCTTCAACAGACTCTCTGAAC 1575
Qy 1501 TGGACCGGGAGGTCTTGCACAGCCCTGCCATCGAGGTCCAGAGTCCACTGCCAGCTGGTT 1560
Db 1576 TGGACCGGGAGGTCTTGCACAGCCCTGCCATCGAGGTCCAGAGTCCACTGCCAGCTGGTT 1635
Qy 1561 CGACTCTTTGCACAGGAAATTGAAGAAATTCGAGCCAGACTCACAACAGCTGA 1614
Db 1636 CGACTCTTTGCACAGGAAATTGAAGAAATTCGAGCCAGACTCACAACAGCTGA 1689

RESULT 4
US-10-631-958-16
; Sequence 16, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631.958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 4413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-16

Query Match 100.0%; Score 1614; DB 19; Length 4413;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGGGCGACGGGGCGGGAGCCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC 60
Db 76 ATGGGGGCGACGGGGCGGGAGCCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC 135
Qy 61 TGGCCCGTGAAGCCCTGAGAGCCCGCGGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 120
Db 136 TGGCCCGTGAAGCCCTGAGAGCCCGCGGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 195
Qy 121 GGAGCCGGCCCGCGGGCGGGATGCTGCTGTGCTGTATCTGAGATCATCGCGTT 180
Db 196 GGAGCCGGCCCGCGGGCGGGATGCTGCTGTGCTGTATCTGAGATCATCGCGTT 255
Qy 181 GAGGAAACAGACGTTTACGGGAAACATCAAGCAGTGGAAATGGCAAAATGGAAAG 240
Db 256 GAGGAAACAGACGTTTACGGGAAACATCAAGCAGTGGAAATGGCAAAATGGAAAG 315
Qy 241 CTTTACGCTTTTACAGTTTCACTGTGTAAGAGACGACGACCGCTGGAAGTGGCG 300
Db 316 CTTTACGCTTTTACAGTTTCACTGTGTAAGAGACGACGACCGCTGGAAGTGGCG 375
Qy 301 CAGTGACTTCTGCTGTCCAGAGAGCAGTGTGCTGCTGTGCTGTGCTGACACCTGCGG 360
Db 376 CAGTGACTTCTGCTGTCCAGAGAGCAGTGTGCTGCTGTGCTGCTGCTGCTGCTGCTG 435
Qy 361 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTTACTGGTATTATCAACCCGTTT 420
Db 436 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTTACTGGTATTATCAACCCGTTT 495
Qy 421 GGAGGAAAGGACAAAGCAAGCGGATATATGAAGAAAGTGGCAACCACTGTTTACCTTA 480
```

```
Db 496 GGAGGAAAGGACAAGGCGGATATATGAAAGAAAGTGGCACCACTGTTTCACTTA 555
Qy 481 GCCTCATCACCACTGACATCATCTGTTACTGAACATGCTTAATCAGGCCAAAGAGACTCTG 540
Db 556 GCCTCCATCACCACCTGACATCATCTGTTACTGAACATGCTTAATCAGGCCAAAGGAGACTCTG 615
Qy 541 TATGAGATTACATAGACAAATACAGCGCATCTGTCGCGGAGATGGTATGTTT 600
Db 616 TATGAGATTAAATAGACAAATACAGCGCATCTGTCGCGGAGATGGTATGTTT 675
Qy 601 AGCAGAGTGTCTCAGCGTCTGATTGGGAGGACGACAGAGAGCGCGGGTCCAGCAGAAC 660
Db 676 AGCAGAGTGTCTCAGCGTCTGATTGGGAGGACGACAGAGAGCGCGGGTCCAGCAGAAC 735
Qy 661 CACCCCGGGCTGTGCTGTGCTCCAGTAGCTCTCCGATGCGGATTTGGAATCATTTCCCGCAGGGTCA 720
Db 736 CACCCCGGGCTGTGCTGTGCTCCAGTAGCTCTCCGATTTGGAATCATTTCCCGCAGGGTCA 795
Qy 721 AGGACTCGTGTGTTACTCCACCGTGGGACGACGAGGACGACAGAACCTCGGCGTGCAT 780
Db 796 AGGACTCGTGTGTTACTCCACCGTGGGACGACGAGGACGACAGAACCTCGGCGTGCAT 855
Qy 781 ATCGTGTGTTGGGACTCGCTGCGCCATGATGTCTCAGTCCACACACACAGACACTC 840
Db 856 ATCGTGTGTTGGGACTCGCTGCGCCATGATGTCTCAGTCCACACACACAGACACTC 915
Qy 841 TTTCGCTACTCGTGTCTGCTGGGCTACCGGCTTCTACGGGACATCATCAAGGACAGT 900
Db 916 TTTCGCTACTCGTGTCTGCTGGGCTACCGGCTTCTACGGGACATCATCAAGGACAGT 975
Qy 901 GAGAGAAACGTTGGTCTTCCAGATACGACTTTTCAAGTTTAAAGACCTTCTC 960
Db 976 GAGAGAAACGTTGGTCTTCCAGATACGACTTTTCAAGTTTAAAGACCTTCTC 1035
Qy 961 TCCACACTGTATGAAAGGACAGTGTCTTCTCCCTGCAACACACAGTGGGATCT 1020
Db 1036 TCCACACTGTATGAAAGGACAGTGTCTTCTCCCTGCAACACACAGTGGGATCT 1095
Qy 1021 CCAAGGGATAGGAAGCCCTGCGGCGAGGATGCTTTGTTTGGAGCAAGCAAGCAGCAG 1080
Db 1096 CCAAGGGATAGGAAGCCCTGCGGCGAGGATGCTTTGTTTGGAGCAAGCAAGCAGCAG 1155
Qy 1081 CTGGAGGAGGACAGAGAAAGCACTGTATGTTTGGAAAGTTCGGAGGACGTGGAGAG 1140
Db 1156 CTGGAGGAGGACAGAGAAAGCACTGTATGTTTGGAAAGTTCGGAGGACGTGGAGAG 1215
Qy 1141 TGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTGTCTTGT 1200
Db 1216 TGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTGTCTTGT 1275
Qy 1201 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGCGGGTCTTTGACCTC 1260
Db 1276 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGCGGGTCTTTGACCTC 1335
Qy 1261 ATCTCTATCCGGAATGTCTCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAAC 1320
Db 1336 ATCTCTATCCGGAATGTCTCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAAC 1395
Qy 1321 CAGCAGGACCACTTTGACTTCACTTTTGTGAAAGTTTATCGCGTCAAGAAATTTCCAGTTT 1380
Db 1396 CAGCAGGACCACTTTGACTTCACTTTTGTGAAAGTTTATCGCGTCAAGAAATTTCCAGTTT 1455
Qy 1381 AGCTCGAAGCACAATGGAGGATGAGGACAGCGACTTCAAGGAGGGGGGGAAGAGCGTTT 1440
Db 1456 AGCTCGAAGCACAATGGAGGATGAGGACAGCGACTTCAAGGAGGGGGGGAAGAGCGTTT 1515
Qy 1441 GGGGCAATTTGGAGAGCACCCTCTCTGCTGCTGCAAGCTCTCAACAGCTCTCTGGAAC 1500
Db 1516 GGGGCAATTTGGAGAGCACCCTCTCTGCTGCTGCAAGCTCTCTCAACAGCTCTCTGGAAC 1575
Qy 1501 TCGCAGGGGAGGTTCTGACAGCCCTGCGCATCGAGGTCAGAGTCCACTGCCAGCTGGTT 1560
```

```
Db 1576 TGCAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGCTGGTT 1635
Qy 1561 CGACTCTTTTGCAGGGAATTTGAAGAGAAATCCGAAGCCAGACTCACAGCTGA 1614
Db 1636 CGACTCTTTTGCAGGGAATTTGAAGAGAAATCCGAAGCCAGACTCACAGCTGA 1689

RESULT 5
US-10-618-941-55
; Sequence 55, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 55
; LENGTH: 4429
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-618-941-55
```

```
Query Match 100.0%; Score 1614; DB 19; Length 4429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGGGCGACGCGGGCGGCGAGCCGCTGCATCCCGTGTGTGGTGAAGCAGCAGCGC 60
Db 92 ATGCGGGCGACGCGGGCGGCGAGCCGCTGCAATCCGTTGCTGAGTGAAGCAGCAGCGC 151
Qy 61 TCGCGCTGTGACCTTGAAGCCCGCGGGGCTCTGTGCGCTGTGTGCGAGCCCGGGGCCC 120
Db 152 TCGCGCTGTGACCTTGAAGCCCGCGGGGCTCTGTGCGCTGTGTGCGAGCCCGGGGCCC 211
Qy 121 GGAGCCGCGCCCGCGCGGCGGATGCTGTCTGTGCTGTATCTGAGATCATCCCGCTT 180
Db 212 GGAGCCGCGCCCGCGCGGCGGATGCTGTCTGTGCTGTATCTGAGATCATCCCGCTT 271
Qy 181 GAGGAAACAGACGTTTCAAGGAAAATCATCAAGGCAAGTGGAAAATGCGAGAAAATGGAAGA 240
Db 272 GAGGAAACAGACGTTTCAAGGAAAATCATCAAGGCAAGTGGAAAATGCGAGAAAATGGAAGA 331
Qy 241 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCACCGCTGGAAGTGGGCG 300
Db 332 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCACCGCTGGAAGTGGGCG 391
Qy 301 CAGGTGACTTTTCTGTGTCCAGAGAGCAGCTGTGTCTGTGCTGTATCTGAGATCATCCCGCTT 360
Db 392 CAGGTGACTTTTCTGTGTCCAGAGAGCAGCTGTGTCTGTGCTGTATCTGAGATCATCCCGCTT 451
Qy 361 GAGATGTGGAGAGCTGACGTCAGAACCAAGCAATTTTACTGGTATTTTCAACCCGTTT 420
Db 452 GAGATGTGGAGAGCTGACGTCAGAACCAAGCAATTTTACTGGTATTTTCAACCCGTTT 511
Qy 421 GGAGAAAGAGACAAAGGACCGGATATATGAAAGAAAAGTGGACCACTGTTTCACTTCA 480
Db 512 GGAGAAAGAGACAAAGGACCGGATATATGAAAGAAAAGTGGACCACTGTTTCACTTCA 571
Qy 481 GCCTCCATCACCCTGACATCATCTGTTACTCAACATGCTTAATCAGGCCAAAGAGACTCTG 540
Db 572 GCCTCCATCACCCTGACATCATCTGTTACTGAAATGCTTAATCAGGCCAAAGAGACTCTG 631
Qy 541 TATGAGATTAAATAGACAAATACAGCGCATCTGTGTGTGCGGAGATGGTATGTTT 600
Db 632 TATGAGATTAAATAGACAAATACAGCGCATCTGTGTGTGCGGAGATGGTATGTTT 691
```

```
Qy 601 AGCGAGTGTGTCACGGTCTGATTGGGAGGACGAGAGCGCGGGGTGCAACGAAAC 660
Db 692 AGCGAGTGTGTCACGGTCTGATTGGGAGGACGAGAGCGCGGGGTGCAACGAAAC 751
Qy 661 CACCCCGGGGTGTGTCGGTCCCGATAGCTCTCGGATTTGGAATCATTTCCCGAGGGTCA 720
Db 752 CACCCCGGGGTGTGTCGGTCCCGATAGCTCTCGGATTTGGAATCATTTCCCGAGGGTCA 811
Qy 721 ACGGACTGCGTGTGTTACTCCACCGTGGGACACGAGCGAGCAAACTCGCGGCTGCAT 780
Db 812 ACGGACTGCGTGTGTTACTCCACCGTGGGACACGAGCGAGCAAACTCGCGGCTGCAT 871
Qy 781 ATCGTTGTTGGGACTCGCTCGCCATGATGTGCTCAGTCCACCAACAGCACACTC 840
Db 872 ATCGTTGTTGGGACTCGCTCGCCATGATGTGCTCAGTCCACCAACAGCACACTC 931
Qy 841 CTTGCTACTCGTGTCTCTGCTGGGCTACCGGCTTTCAGGGGACATCATCAAGGACAGT 900
Db 932 CTTGCTACTCGTGTCTCTGCTGGGCTACCGGCTTTCAGGGGACATCATCAAGGACAGT 991
Qy 901 GAGAAGAACCGTGGTGTCTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC 960
Db 992 GAGAAGAACCGTGGTGTCTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC 1051
Qy 961 TCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACAGCGTGGATCT 1020
Db 1052 TCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACAGCGTGGATCT 1111
Qy 1021 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGGAG 1080
Db 1112 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGGAG 1171
Qy 1081 CTGGAGGAGGACGAGAAAGACATGATGTTTGAAGCTGCGGAGGACGTGGAGGAG 1140
Db 1172 CTGGAGGAGGACGAGAAAGACATGATGTTTGAAGCTGCGGAGGACGTGGAGGAG 1231
Qy 1141 TGGCAAGTCGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTGTGCTTGT 1200
Db 1232 TGGCAAGTCGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTGTGCTTGT 1291
Qy 1201 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACATTTGGGAGAGCGGTCTTTCGACCTC 1260
Db 1292 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACATTTGGGAGAGCGGTCTTTCGACCTC 1351
Qy 1261 ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAAC 1320
Db 1352 ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAAC 1411
Qy 1321 CAGCAGGACCAAGTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
Db 1412 CAGCAGGACCAAGTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1471
Qy 1381 ACCTCGAAGCATTGAGGATGAGGACAGGACCTCAAGGAGGGGGGGAAGAGGCTTT 1440
Db 1472 ACCTCGAAGCATTGAGGATGAGGACAGGACCTCAAGGAGGGGGGGAAGAGGCTTT 1531
Qy 1441 GGGCACATTTTCAGCAGCACCCCTCTGCTGCTGTCACCGTCTCCACAGCTCTCGGAAC 1500
Db 1532 GGGCACATTTTCAGCAGCACCCCTCTGCTGCTGTCACCGTCTCCACAGCTCTCGGAAC 1591
Qy 1501 TCGCAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCCAGTGGTT 1560
Db 1592 TCGCAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCCAGTGGTT 1651
Qy 1561 CGACTCTTTGACGAGGAATTTGAAGAAATTCGAAGCCAGACTCACAGCTGA 1614
Db 1652 CGACTCTTTGACGAGGAATTTGAAGAAATTCGAAGCCAGACTCACAGCTGA 1705
```

```
RESULT 6
US-10-120-988-148
; Sequence 148, Application US/10120988
; Publication No. US20030219745A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt-FL_genes Version 2.0
; SEQ ID NO 148
; LENGTH: 4432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129)..(2817)
; US-10-120-988-148

Query Match 99.9%; Score 1612.4; DB 17; Length 4432;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1613; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 ATGGGGGCGACGGGGGCGGGAGCCGCTGCAATCCGTGCTGTGGGTGAAGCAGCAGGC 60
Db 1204 ATGGGGGCGACGGGGGCGGGAGCCGCTGCAATCCGTGCTGTGGGTGAAGCAGCAGGC 1263
Qy 61 TGGCCCTGAGCTTGGAGCCGCGGGCTCTGCTGCTGTATCTGAGATCATCGCGTT 180
Db 1264 TGGCCCTGAGCTTGGAGCCGCGGGCTCTGCTGCTGTATCTGAGATCATCGCGTT 1323
Qy 121 GGAGCCGCGCCCGCGGGGATGCTGCTGCTGTATCTGAGATCATCGCGTT 180
Db 1324 GGAGCCGCGCCCGCGGGGATGCTGCTGCTGTATCTGAGATCATCGCGTT 1383
Qy 181 GAGAAACAGACGTTTCACGGGAAACATCAAGGACAGTGGAAATGGCAAAATGGAAG 240
Db 1384 GAGAAACAGACGTTTCACGGGAAACATCAAGGACAGTGGAAATGGCAAAATGGAAG 1443
Qy 241 CCTTACCGTTTACAGTTCACTGTAAAGAGACAGCGGACCGCTGGAAGTGGCG 300
Db 1444 CCTTACCGTTTACAGTTCACTGTAAAGAGACAGCGGACCGCTGGAAGTGGCG 1503
Qy 301 CAGGTGACTTTTCTGGTGTCCAGAGGACGCTGTGTCACTTGTGGTGCAGACCTCGCG 360
Db 1504 CAGGTGACTTTTCTGGTGTCCAGAGGACGCTGTGTCACTTGTGGTGCAGACCTCGCG 1563
Qy 361 GAGATGTGGAGAGCTGACGTCCAGACCAAGCAATTTACTGGTATTATCAACCCGTTT 420
Db 1564 GAGATGTGGAGAGCTGACGTCCAGACCAAGCAATTTACTGGTATTATCAACCCGTTT 1623
Qy 421 GGAGAAAGGACAAAGCAAGCGGATATATGAAAGAAAGTGCCACCACTGTTCAACCTTA 480
Db 1624 GGAGAAAGGACAAAGCAAGCGGATATATGAAAGAAAGTGCCACCACTGTTCAACCTTA 1683
Qy 481 GCCTTCCACTCACTGACATCATCGTTTACTGAAACATGCTAATCAGGCCAAGGAGACTCTG 540
Db 1684 GCCTTCCACTCACTGACATCATCGTTTACTGAAACATGCTAATCAGGCCAAGGAGACTCTG 1743
Qy 541 TATGAGATTAACTAGACAAATACGAGGCACTGCTGTGTGCGCGGAGATGGTATGTTTC 600
Db 1744 TATGAGATTAACTAGACAAATACGAGGCACTGCTGTGTGCGCGGAGATGGTATGTTTC 1803
Qy 601 AGCGAGGTGTGTCACGGTCTGATTGGGAGGACGAGAGGAGCGCGGGGTCCAGCAGAAC 660
Db 1804 AGCGAGGTGTGTCACGGTCTGATTGGGAGGACGAGAGGAGCGCGGGGTCCAGCAGAAC 1863
```


Qy	661	CACCC	CGGG	CTGTG	CTGGT	CCCC	CAGT	AGCTC	CGGAT	TGGAAT	CAAT	TCCG	CAGGGT	CA	720				
Db	1864	CACCC	CGGG	CTGTG	CTGGT	CCCC	CAGT	AGCTC	CGGAT	TGGAAT	CAAT	TCCG	CAGGGT	CA	1923				
Qy	721	ACG	ACTC	GC	GTG	TTACT	CTC	CAC	CGTGGG	CAC	AGC	GAC	GAGAA	CCTCGG	CGCTGCAT	780			
Db	1924	ACG	ACTC	GC	GTG	TTACT	CTC	CAC	CGTGGG	CAC	AGC	GAC	GAGAA	CCTCGG	CGCTGCAT	1983			
Qy	781	ATC	GTTT	TG	TGGG	GA	CTCG	TG	CCAT	GGAT	GTG	TCT	CAGT	CCAC	CAAC	CAGC	CACTC	840	
Db	1984	ATC	GTTT	TG	TGGG	GA	CTCG	TG	CCAT	GGAT	GTG	TCT	CAGT	CCAC	CAAC	CAGC	CACTC	2043	
Qy	841	CTT	CGT	ACT	CCG	TG	TCC	TG	CTG	CTGG	CTA	CGG	CTT	CTA	CGG	GGA	CAT	CA	900
Db	2044	CTT	CGT	ACT	CCG	TG	TCC	TG	CTG	CTGG	CTA	CGG	CTT	CTA	CGG	GGA	CAT	CA	2103
Qy	901	GAG	AAGAA	ACG	TG	TGGG	TCTT	G	CCAG	ATAC	CAC	ATTTT	CAG	GT	TAA	AG	ACCTT	CCTC	960
Db	2104	GAG	AAGAA	ACG	TG	TGGG	TCTT	G	CCAG	ATAC	CAC	ATTTT	CAG	GT	TAA	AG	ACCTT	CCTC	2163
Qy	961	TCC	CAC	CA	CTG	TAT	G	AAGG	GAC	AGT	GT	TCTT	CTC	CTG	CA	CA	CA	CA	1020
Db	2164	TCC	CAC	CA	CTG	TAT	G	AAGG	GAC	AGT	GT	TCTT	CTC	CTG	CA	CA	CA	CA	2223
Qy	1021	CCA	AGG	AT	AGG	AA	CCCTG	CCG	GGC	AG	ATG	CTT	TTG	TG	CA	GGC	AAA	CA	1080
Db	2224	CCA	AGG	AT	AGG	AA	CCCTG	CCG	GGC	AG	ATG	CTT	TTG	TG	CA	GGC	AAA	CA	2283
Qy	1081	CTG	GAG	GAG	GAG	CAG	AG	AA	AG	CA	CTG	TAT	GGT	TG	GA	CTG	CGG	AGG	1140
Db	2284	CTG	GAG	GAG	GAG	CAG	AG	AA	AG	CA	CTG	TAT	GGT	TG	GA	CTG	CGG	AGG	2343
Qy	1141	TGG	CA	AGT	CTG	CTG	TGG	GA	AGT	TTT	CTG	GC	CA	AT	CA	AT	GC	CA	1200
Db	2344	TGG	CA	AGT	CTG	CTG	TGG	GA	AGT	TTT	CTG	GC	CA	AT	CA	AT	GC	CA	2403
Qy	1201	CGC	CG	AG	CCCC	CA	GGG	GC	CTC	CCG	GC	TGCC	CA	CTG	G	G	AG	AC	1260
Db	2404	CGC	CG	AG	CCCC	CA	GGG	GC	CTC	CCG	GC	TGCC	CA	CTG	G	G	AG	AC	2463
Qy	1261	ATC	CT	CA	TC	CGG	AA	TG	CT	C	CAG	GT	CA	ATTTT	CT	G	A	G	1320
Db	2464	ATC	CT	CA	TC	CGG	AA	TG	CT	C	CAG	GT	CA	ATTTT	CT	G	A	G	2523
Qy	1321	CAG	CAG	GA	CC	AGT	TTG	A	CTT	CA	CTTTT	TG	TG	AA	GT	TAT	CG	CGT	1380
Db	2524	CAG	CAG	GA	CC	AGT	TTG	A	CTT	CA	CTTTT	TG	TG	AA	GT	TAT	CG	CGT	2583
Qy	1381	ACG	T	CA	AG	CA	CAT	G	GA	ATG	AG	GA	CAG	CGA	CTC	T	A	G	1440
Db	2584	ACG	T	CA	AG	CA	CAT	G	GA	ATG	AG	GA	CAG	CGA	CTC	T	A	G	2643
Qy	1441	GGG	CA	CA	TTT	TG	CAG	AG	CC	AC	CC	CT	CTG	CTG	CA	CC	GT	CT	1500
Db	2644	GGG	CA	CA	TTT	TG	CAG	AG	CC	AC	CC	CT	CTG	CTG	CA	CC	GT	CT	2703
Qy	1501	TG	C	A	C	G	G	G	A	G	T	C	T	G	C	A	C	A	1560
Db	2704	TG	C	A	C	G	G	A	G	T	C	T	G	C	A	C	A	C	2763
Qy	1561	CGA	CT	TTT	TG	CA	CG	AG	AA	TTG	AA	GA	GA	TCC	GA	AG			

RESULT 7

US-101-262-511-39
; Sequence 39, Application US/10262511
; Publication No. J520040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
;

```

1  / APPLICANT: Kekuda, Ramesh
2  / APPLICANT: Ju, Jingfang
3  / APPLICANT: Li, Li
4  / APPLICANT: Guo, Xiaojia (Saaha)
5  / APPLICANT: Patturajan, Meera
6  / APPLICANT: Spytek, Kimberly A.
7  / APPLICANT: Edinger, Shlomit R.
8  / APPLICANT: Ellerman, Karen
9  / APPLICANT: Malvankar, Uriel M.
10 / APPLICANT: Ort, Tatiana
11 / APPLICANT: Gorman, Linda
12 / APPLICANT: Zerhusen, Bryan D.
13 / APPLICANT: Anderson, David W.
14 / APPLICANT: Zhong, Mei
15 / APPLICANT: Catterton, Elina
16 / APPLICANT: Ji, Weizhen
17 / APPLICANT: Miller, Charles E.
18 / APPLICANT: Rastelli, Luca
19 / APPLICANT: Stone, David J.
20 / APPLICANT: Pena, Carol E. A.
21 / APPLICANT: Shenoy, Suresh G.
22 / APPLICANT: Shimkets, Richard A.
23 / APPLICANT: Rothenberg, Mark E.
24 / APPLICANT: Leach, Martin D.
25 / APPLICANT: Agee, Michele L.
26 / APPLICANT: Berghs, Constance
27 / TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
28 / FILE REFERENCE: 21402-462C
29 / CURRENT APPLICATION NUMBER: US/10/262,511
30 / CURRENT FILING DATE: 2003-05-28
31 / PRIOR APPLICATION NUMBER: 60/326,483
32 / PRIOR FILING DATE: 2001-10-02
33 / PRIOR APPLICATION NUMBER: 60/373,815
34 / PRIOR FILING DATE: 2002-04-19
35 / PRIOR APPLICATION NUMBER: 60/327,917
36 / PRIOR FILING DATE: 2001-10-09
37 / PRIOR APPLICATION NUMBER: 60/381,642
38 / PRIOR FILING DATE: 2002-05-17
39 / PRIOR APPLICATION NUMBER: 60/328,029
40 / PRIOR FILING DATE: 2002-10-09
41 / PRIOR APPLICATION NUMBER: 60/381,038
42 / PRIOR FILING DATE: 2002-05-16
43 / PRIOR APPLICATION NUMBER: 60/328,056
44 / PRIOR FILING DATE: 2001-10-09
45 / PRIOR APPLICATION NUMBER: 60/373,260
46 / PRIOR FILING DATE: 2002-04-17
47 / PRIOR APPLICATION NUMBER: 60/373,826
48 / PRIOR FILING DATE: 2002-04-19
49 / PRIOR APPLICATION NUMBER: 60/327,435
50 / PRIOR FILING DATE: 2001-10-05
51 / Remaining Prior Application data removed - See File Wrapper or PALN
52 / NUMBER OF SEQ ID NOS: 439
53 / SOFTWARE: CuraSeqList version 0.1
54 / SEQ ID NO 39
55 / LENGTH: 1740
56 / TYPE: DNA
57 / ORGANISM: Homo sapiens
58 / FEATURE:
59 / NAME/KEY: CDS
60 / LOCATION: (76)..(1686)
61 / US-10-262-511-39

```

Query Match 99.8%; Score 1610.8; DB 18; Length 1740;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 1612; Conservative 0; Mismatches 2; Indels 0; Gaps 0

1	ATGGGGCGGACGGGGGCGGCGAGCCCTCGAATCCCTGCTGTGGGTGAACGACGAGCGC	60
76	ATGGGGCGGACGGGGGCGGCGAGCCCTCGAATCCCTGCTGTGGGTGAACGAGCAGCGC	135
61	TGGCCCGTGAGCCTTGGAGCCCGCGCGGGCTCTGTGTGGCTGTGGCGAGAGCCCGGGGCC	120
136	TGGCCCGTGAGCCTTGGAGCCCGCGCGGGCTCTGTGTGGCTGTGGCGAGAGCCCGGGGCC	195

Qy	121	GGAGCCGGCGCCCCCGCGCGGATGCGCTGCTCTGTGCTGTATCTGAGATCATCGCGGTT	180
Db	196	GGAGCCGGCGCCCCCGGTGTGATGCTGTCTGTGCTGTATCTGAGATCATCGCGGTT	255
Qy	181	GAGGAAACAGACGTTCAACGGGAAACATCAAGGCAGTGGAAAATGGCAGAAAATGGAAAAG	240
Db	256	GAGGAAACAGACGTTCAACGGGAAACATCAAGGCAGTGGAAAATGGCAGAAAATGGAAAAG	315
Qy	241	CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGCGCACCGCTGGAAAGTGGCG	300
Db	316	CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGCGCACCGCTGGAAAGTGGCG	375
Qy	301	CAGGTGACTTTCTGCTGTCCAGAGGACGCTGTGTCACTTGTGTGCTGCAGACCTTCGGG	360
Db	376	CAGGTGACTTTCTGCTGTCCAGAGGACGCTGTGTCACTTGTGTGCTGCAGACCTTCGGG	435
Qy	361	GAGATGCTGGAGAAGCTGACGTGCAGACCAAAAGCATTTACTGTGTATTTATCAACCCGTTT	420
Db	436	GAGATGCTGGAGAAGCTGACGTGCAGACCAAAAGCATTTACTGTGTATTTATCAACCCGTTT	495
Qy	421	GGAGGAAAAGGACAAGGCAAGCGGATATATGAAGAAAAGTGGCACCACTGTTTCACCTTA	480
Db	496	GGAGGAAAAGGACAAGGCAAGCGGATATATGAAGAAAAGTGGCACCACTGTTTCACCTTA	555
Qy	481	GCCTCATACCACTGACATCATCGTTACTTGAACATATGCTTAATCAGGCCAAGGAGACTCTG	540
Db	556	GCCTCCATCACCATGACATCATCGTTACTTGAACATATGCTTAATCAGGCCAAGGAGACTCTG	615
Qy	541	TATGAGATTAAACATAGACAATAACGCGCATCGTCTGCTCGCGCGAGATGTTATGTTTC	600
Db	616	TATGAGATTAAACATAGACAATAACGCGCATCGTCTGCTCGCGCGAGATGTTATGTTTC	675
Qy	601	AGCAGGTGCTGCACGGTCTGATTTGGAGGACGACAGAGAGCGCCGGGTCGACCAAGAAC	660
Db	676	AGCAGGTGCTGCACGGTCTGATTTGGAGGACGACAGAGAGCGCCGGGTCGACCAAGAAC	735
Qy	661	CACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGGATTGGATTCATTCGCGCAGAGTCA	720
Db	736	CACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGGATTGGATTCATTCGCGCAGAGTCA	795
Qy	721	ACGACTGCGTGTGTTACTCCACCGTGGGCACACAGCAGCAGAAAACCTCGCGCGCTGCAT	780
Db	796	ACGACTGCGTGTGTTACTCCA CGTGGGCACACAGCAGCAGAAAACCTCGCGCGCTGCAT	855
Qy	781	ATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACACAGCACTC	840
Db	856	ATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACACAGCACTC	915
Qy	841	CTTCGCTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACGGGCATCATCAAGGACAGT	900
Db	916	CTTCGCTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACGGGCATCATCAAGGACAGT	975
Qy	901	GAGAAAGAAACGGTGGTGGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC	960
Db	976	GAGAAAGAAACGGTGGTGGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC	1035
Qy	961	TCCCAACACTGTATGAAGGAGACAGTGTCTCTTCCCTGCAACAAACACAGCGTGGGATCT	1020
Db	1036	TCCCAACACTGTATGAAGGAGACAGTGTCTCTTCCCTGCAACAAACACAGCGTGGGATCT	1095
Qy	1021	CCAGGAGATAGGAGCCCTCCGGCAGGATGCTTTGTTTTCAGGCAAGACGACGACGAG	1080
Db	1096	CCAAGGATAGGAGCCCTCCGGCAGGATGCTTTGTTTTCAGGCAAGACGACGACGAG	1155
Qy	1081	CTGAGGAGGAGCAGAAAGAAAGCACTGTATGTGTTTGGAAAGCTCGGAGGACGTGGAGG	1140
Db	1156	CTGAGGAGGAGCAGAAAGAAAGCACTGTATGTGTTTGGAAAGCTCGGAGGACGTGGAGG	1215
Qy	1141	TGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCCAAAACATGTCTGTGCTGTGT	1200
Db	1216	TGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCCAAAACATGTCTGTGCTGTGT	1275

Qy	1201	GGCGGAGCCCCAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGCTCTTCTGACCTC	1261
Db	1276	GGCGGAGCCCCAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	1335
Qy	1261	ATCCTCATCCGGAATGCTCCAGGTTCAAATTTTCTGAGATTCTCATCAGGCACACCAAC	1320
Db	1336	ATCCTCATCCGGAATGCTCCAGGTTCAAATTTTCTGAGATTCTCATCAGGCACACCAAC	1395
Qy	1321	CAGCAGGACAGTTTGACTTCACTTTTGTGTAAGTTTATCGCGTCAAGAAATTCAGTTT	1380
Db	1396	CAGCAGGACAGTTTGACTTCACTTTTGTGTAAGTTTATCGCGTCAAGAAATTCAGTTT	1455
Qy	1381	ACGTCGAAGCACATGGAGGATGAGGACAGCGACTCAAGCAGGGGGGGAAGAGCGCTTT	1440
Db	1456	ACGTCGAAGCACATGGAGGATGAGGACAGCGACTCAAGCAGGGGGGGAAGAGCGCTTT	1515
Qy	1441	GGGCACATTTGCAGCAGCCACCCCTCTGCTGTGTCACCGCTCTCAAACAGCTCTCGAAC	1500
Db	1516	GGGCACATTTGCAGCAGCCACCCCTCTGCTGTGTCACCGCTCTCAAACAGCTCTCGAAC	1575
Qy	1501	TGCGACGGGAGTCTTCGACAGCCCTGCCATCGAGGTCAGAGTCCCACTGCCAGCTGTT	1560
Db	1576	TGCGACGGGAGTCTTCGACAGCCCTGCCATCGAGGTCAGAGTCCCACTGCCAGCTGTT	1635
Qy	1561	CGACTCTTTGCACGAGGAATTGAAGAAATCCGAAGCCAGACTCACACAGCTGA	1614
Db	1636	CGACTCTTTGCACGAGGAATTGAAGAAATCCGAAGCCAGACTCACACAGCTGA	1689
RESULT 8			
US-10-315-597A-1			
; Sequence 1, Application US/10315597A			
; Publication No. US20030162206A1			
; GENERAL INFORMATION:			
; APPLICANT: Kono, Masako			
; APPLICANT: Sugiura, Keita			
; APPLICANT: Kohama, Takafumi			
; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It			
; FILE REFERENCE: 02658CIP/HG			
; CURRENT APPLICATION NUMBER: US/10/315,597A			
; CURRENT FILING DATE: 2002-12-10			
; PRIOR APPLICATION NUMBER: JP 2000-178039			
; PRIOR FILING DATE: 2000-06-14			
; NUMBER OF SEQ ID NOS: 4			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 1			
; LENGTH: 4463			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: unsure			
; LOCATION: 3371			
; OTHER INFORMATION:			
US-10-315-597A-1			
Query Match 99.6%; Score 1607.6; DB 16; Length 4463;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1610; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
Qy	1	ATGGGGGCGACGGGGCGGGAGCCGCTGCAATCCGTGCTGTGGGTGAAGCAGCAGCGC	60
Db	124	ATGGGGGCGACGGGGCGGGAGCCGCTGCAATCCGTGCTGTGGGTGAAGCAGCAGCGC	183
Qy	61	TGCGCCCGTGAGCCTTGAGCCCGCGCGGCTCTGCTGCGCTGTGGTGGAGGCCCGGCC	120
Db	184	TGCGCCCGTGAGCCTTGAGCCCGCGCGGCTCTGCTGCGCTGTGGTGGAGGCCCGGCC	243
Qy	121	GGACCGGCGCCCCGGGGCGGATGCTGCTGTGCTGTGCTGTGATCTGAGATCATCGCGTT	180
Db	244	GGACCGGCGCCCCGGGGCGGATGCTGCTGTGCTGTGCTGTGATCTGAGATCATCGCGTT	303
Qy	181	GAGGAAAACAGACGTTTCACGGGAAAACATCAAGGCAGTGGAAAATGGCAGAAAATGGA	240

304	Db	GAGGAAACAGACGTTCCAGCGGAAACATCAAGCGCAGTGGAAATATGGCAGAAAAATGGCAAAATGGAAAG	363
241	Qy	CTTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGACGGCACCGCTGGAAGTGGCG	300
364	Db	CTTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGACGACGCGCACCGCTGGAGTGGCG	423
301	Qy	CAGGTGACTTTCGTGGTGTCCAGAGGAGCAGCTGTGTCACTTTGTGGCTGCGACACCCTCGG	360
424	Db	CAGGTGACTTTCGTGGTGTCCAGAGGAGCAGCTGTGTCACTTTGTGGCTGCGACACCCTCGG	483
361	Qy	GAGATGCTGGAGAAGCTGACGTGCAGACCAAGCATTTACTGTGTATTTATCAACCCGTTT	420
484	Db	GAGATGCTGGAGAAGCTGACGTGCAGACCAAGCATTTACTGTGTATTTATCAACCCGTTT	543
421	Qy	GGAGGAAAAGGACAAAGCGGATATATGAAAGAAAAGTGGCAACAATGTTTCAACCTTA	480
544	Db	GGAGGAAAAGGACAAAGCGGATATATGAAAGAAAAGTGGCAACAATGTTTCAACCTTA	603
481	Qy	GCCTCCATCACCATGACATCATCGTTACTGAAATGCTTAATCAGGCCAAGGAGACTCTG	540
604	Db	GCCTCCATCACCATGACATCATCGTTACTGAAATGCTTAATCAGGCCAAGGAGACTCTG	663
541	Qy	TATGAGATTAAATAGACAAATACGAGGCAATCGTCTGTGTGCGCGGAGATGGTATGTC	600
664	Db	TATGAGATTAAATAGACAAATACGAGGCAATCGTCTGTGTGCGCGGAGATGGTATGTC	723
601	Qy	AGCAGGTGCTGCAGCGTCTGATTTGGGAGACGACGAGGAGCGCGCGGTTCGACAGAAC	660
724	Db	AGCAGGTGCTGCAGCGTCTGATTTGGGAGACGACGAGGAGCGCGCGGTTCGACAGAAC	783
661	Qy	CACCCCGGGCTGTGCTGGTCCCAGTAGCCTCCGGATTGGAATCAATCCCGCAGGTCCA	720
784	Db	CACCCCGGGCTGTGCTGGTCCCAGTAGCCTCCGGATTGGAATCAATCCCGCAGGTCCA	843
721	Qy	ACGSACTCCGTGTGTTACTCCACCGTGGGCACAGCGACGCGAGAAACCTCGGCGCTGCAT	780
844	Db	ACGSACTCCGTGTGTTACTCCACCGTGGGCACAGCGACGCGAGAAACCTCGGCGCTGCAT	903
781	Qy	ATCGTTGTTGGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAACAGCACACTC	840
904	Db	ATCGTTGTTGGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAACAGCACACTC	963
841	Qy	CTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGT	900
964	Db	CTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGT	1023
901	Qy	GAGAGAAAACGGTGGTGGGTCTGCGACATACGACTTTTCAGGTTTAAAGACCTTCCTC	960
1024	Db	GAGAGAAAACGGTGGTGGGTCTTTCGAGATACGACTTTTCAGGTTTAAAGACCTTCCTC	1083
961	Qy	TCCCAACCTGCTATGAAGGACAGTGTCTTCTCTCCCTGCAACAACAACGTTGGGATCT	1020
1084	Db	TCCCAACCTGCTATGAAGGACAGTGTCTTCTCTCCCTGCAACAACAACGTTGGGATCT	1143
1021	Qy	CCAAGGATAGGAAGCCCTGCGCGGACAGTGTCTTCTTTCGAGGCAAAAGCAAGCAGAG	1080
1144	Db	CCAAGGATAGGAAGCCCTGCGCGGACAGTGTCTTCTTTCGAGGCAAAAGCAAGCAGAG	1203
1081	Qy	CTGAGAGAGAGCAAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAGGAGCTGTGGAGG	1140
1204	Db	CTGAGAGAGAGCAAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAGGAGCTGTGGAGG	1263
1141	Qy	TGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCACAACAATGCTCTGTGCTTGT	1200
1264	Db	TGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCACAACAATGCTCTGTGCTTGT	1323
1201	Qy	CGCCGAGCCCCAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	1260
1324	Db	CGCCGAGCCCCAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	1383
1261	Qy	ATCTCTCATCCGGAATGCTCAGGTTCAAATTTCTGTAGATTTCTCATCAGGCAACCAAC	1320
1384	Db	ATCTCTCATCCGGAATGCTCAGGTTCAAATTTCTGTAGATTTCTCATCAGGCAACCAAC	1443

Qy	1321	CAGCAGACACAGTTTGACTTTCATCTTTTGTGTGAAGTTTATCGCGTCGAAGAAATTCACAGTTT	1380
Db	1444	CAGCAGACACAGTTTGACTTTCATCTTTTGTGTGAAGTTTATCGCGTCGAAGAAATTCACAGTTT	1503
Qy	1381	ACGTTCGAAGCACATGGAGGATGAGGACAGCGACCTTCAAGGAGGGGGGGAAGAGCCCTTT	1440
Db	1504	ACGTTCGAAGCACATGGAGGATGAGGACAGCGACCTTCAAGGAGGGGGGGAAGAGCCCTTT	1563
Qy	1441	GGGCACATTTGCAGCAGCCACCCCTCTGTGTGTGACCCGTCTTCAACAGCTCTCTGGAAC	1500
Db	1564	GGGCACATTTGCAGCAGCCACCCCTCTGTGTGTGACCCGTCTTCAACAGCTCTCTGGAAC	1623
Qy	1501	TGGCAGGGGAGGTCTGTGCACAGCCCTGCCATCGAGGTGAGAGTCCACTGCCAGCTGGTT	1560
Db	1624	TGGATGGGAGGTCTGTGCACAGCCCTGCCATCGAGGTGAGAGTCCACTGCCAGCTGGTT	1683
Qy	1561	CGACTCTTTGCCAGAGGAATTGAAGAGAAATCGAAGCCAGACTCACACAGCTGA	1614
Db	1684	CGACTCTTTGCCAGAGGAATTGAAGAGAAATCGAAGCCAGACTCACACAGCTGA	1737

RESULT 9

RESUL-9
 US-09-784-810A-5
 ; Sequence 5, Application US/09784810A
 ; Patent No. US20020082203A1
 ; GENERAL INFORMATION:
 ; APPLICANT: RASTELLI, LUCA
 ; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
 ; TITLE OF INVENTION: SAME
 ; FILE REFERENCE: 10716-08
 ; CURRENT APPLICATION NUMBER: US/09/784,810A
 ; CURRENT FILING DATE: 2001-02-14
 ; PRIOR APPLICATION NUMBER: 60/182,360
 ; PRIOR FILING DATE: 2000-02-14
 ; PRIOR APPLICATION NUMBER: 60/191,261
 ; PRIOR FILING DATE: 2000-03-22
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 1840
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-784-810A-5

Query Match	89.6%	Score 1446.4	DB 9	Length 1840
Best Local Similarity	97.8%	Pred. No. 0		
Matches 1490	Conservative 0	Mismatches 1	Indels 33	Gaps 1
Qy	124	GC CGCGCGCCCCGCGCGGATGCCTGCTCTGTGCGCTGTATCTGAGATCATCGCCGTTGAG	183	
Db	1	GCGCGCGCCCCGCGCGGATGCCTGCTCTGTGCGCTGTATCTGAGATCATCGCCGTTGAG	60	
Qy	184	GA AAACAGACGTTTCA CGGGAAAACATCAAGGCAGTGGAAAAATGGCAGAAAAATGGA AAGCCT	243	
Db	61	GA AACAGACGTTTCA CGGGAAAACATCAAGGCAGTGGAAAAATGGCAGAAAAATGGA AAGCCT	120	
Qy	244	TACGCTTTTACAGTTCACTGTGTPAAAGAGACGACGACCGCTGGGAAGTGGGCGCAG	303	
Db	121	TACGCTTTTACAGTTCACTGTGTPAAAGAGACGACGACCGCTGGGAAGTGGGCGCAG	180	
Qy	304	GTACATTTCTGGTGTCAGAGGGAGCCTGTGTCACTTGTGCTGCAGACCCCTGCGGGAG	363	
Db	181	GTACATTTCTGGTGTCAGAGGGAGCCTGTGTCACTTGTGCTGCAGACCCCTGCGGGAG	240	
Qy	364	ATGCTGGAGAAGCTGACGCTCCAGACCAAAGCACTTTACTGGTATTTTCAACCCGTTTGA	423	
Db	241	ATGCTGGAGAAGCTGACGCTCCAGACCAAAGCACTTTACTGGTATTTTCAACCCGTTTGA	300	
Qy	424	GG AAAAGGACAGGCAAGCGGATATATGAAAGAAAAGTGGCAACCACTGTTCACTTAGCC	483	
Db	301	GG AAAAGGACAGGCAAGCGGATATATGAAAGAAAAGTGGCAACCACTGTTCACTTAGCC	360	

Qy 484 TCCATCACCCTGACATCATCG-----TTACT 510
Db 361 TCCATCACCCTGACATCATCGGTAAACAAATTCATATGTTAACTATGTAGAAGTAATTA 420
Qy 511 GAAATGCTAATCAGGCCAAGAGACTCTGTATGAGATTACATAGACAAATACGACGGC 570
Db 421 GAAATGCTAATCAGGCCAAGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGC 480
Qy 571 ATCGTCTGTGCGGGAGATGGTATGTTTCAGCGAGGTGTCACGGGTCTGATTGGGAGG 630
Db 481 ATCGTCTGTGCGGGAGATGGTATGTTTCAGCGAGGTGTCACGGGTCTGATTGGGAGG 540
Qy 631 ACGCAGAGGAGCGCGGGGTCCAGCAGAAACACCCCGGGCTGTCTGGTCCCGCAGTAGC 690
Db 541 ACGCAGAGGAGCGCGGGGTCCAGCAGAAACACCCCGGGCTGTCTGGTCCCGCAGTAGC 600
Qy 691 CTCGGATTGAATCATTTCCCGCAGGGTCAACGAGACTGCGTGTGTTACTCCACCGTGGGC 750
Db 601 CTCGGATTGAATCATTTCCCGCAGGGTCAACGAGACTGCGTGTGTTACTCCACCGTGGGC 660
Qy 751 ACCAGCAGCAGAAACCTCGGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGAT 810
Db 661 ACCAGCAGCAGAAACCTCGGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGAT 720
Qy 811 GTGTCTCAGTCCACCAACAGCAGACACTCCTTCGCTACTCCGTGTCCTGCTGGGCTAC 870
Db 721 GTGTCTCAGTCCACCAACAGCAGACACTCCTTCGCTACTCCGTGTCCTGCTGGGCTAC 780
Qy 871 GGCTTCTACGGGGACATCATCAGGACAGTGAAGAAACGGTGTGGGTCTTGCACGA 930
Db 781 GGCTTCTACGGGGACATCATCAGGACAGTGAAGAAACGGTGTGGGTCTTGCACGA 840
Qy 931 TAGCACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTCTATGAAGGACAGTGTCC 990
Db 841 TAGCACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTCTATGAAGGACAGTGTCC 900
Qy 991 TTCTCTCCTGCACAAACACGCGTGGATCTCCAAAGGATAGAAACCCCTGCGGGCAGGA 1050
Db 901 TTCTCTCCTGCACAAACACGCGTGGATCTCCAAAGGATAGAAACCCCTGCGGGCAGGA 960
Qy 1051 TGCTTTGTTTCAGGCAAAAGCAGCAGCTGAGGAGGAGCAGAGAAAGCACTGTAT 1110
Db 961 TGCTTTGTTTCAGGCAAAAGCAGCAGCTGAGGAGGAGCAGAGAAAGCACTGTAT 1020
Qy 1111 GGTTTGGAAGCTGCGAGGAGCTGAGGAGTGGCAAGTCTGTGGGAAAGTTCTTGCC 1170
Db 1021 GGTTTGGAAGCTGCGAGGAGCTGAGGAGTGGCAAGTCTGTGGGAAAGTTCTTGCC 1080
Qy 1171 ATCAATGCCACAAACATGTCCTGTGTTGTGCGCGAGCCCGAGGGGCTCTCCCGGCT 1230
Db 1081 ATCAATGCCACAAACATGTCCTGTGTTGTGCGCGAGCCCGAGGGGCTCTCCCGGCT 1140
Qy 1231 GCCCACTTGGGAGAGCGGTCTTTCAGCTCATCTCATCCGGAATGCTCCAGGTTCAAT 1290
Db 1141 GCCCACTTGGGAGAGCGGTCTTTCAGCTCATCTCATCCGGAATGCTCCAGGTTCAAT 1200
Qy 1291 TTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCAAGTTTGACTTCACCTTTTGT 1350
Db 1201 TTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCAAGTTTGACTTCACCTTTTGT 1260
Qy 1351 GAAAGTTATCGCGTCAAGAAATTCAGTTTACGTCGAAGCAGCATGGAGGATGAGACAGC 1410
Db 1261 GAAAGTTATCGCGTCAAGAAATTCAGTTTACGTCGAAGCAGCATGGAGGATGAGACAGC 1320
Qy 1411 GACCTCAAGAGGGGGGAGAGCGCTTTGGGCAATTTGACAGCAGCCACCCCTCCTGC 1470
Db 1321 GACCTCAAGAGGGGGGAGAGCGCTTTGGGCAATTTGACAGCAGCCACCCCTCCTGC 1380
Qy 1471 TGCTGACCGCTCCAAACAGCTCTCGAACTGCGACGGGAGGTCTCTGACAGCCCTGCC 1530
Db 1381 TGCTGACCGCTCCAAACAGCTCTCGAACTGCGACGGGAGGTCTCTGACAGCCCTGCC 1440
Qy 1531 ATCGAGGTGAGAGTCCACTGCGCAGCTGGTTCGACTCTTTTGACAGGAAATTGAAGAGAAT 1590

Db 1441 ATCAGGTCAGGGTCCACTGCCAGCTGTTGCACTCTTTGCCAGGGAATTGAAGAAAT 1500
Qy 1591 CCGAAGCCAGACTCACACAGCTGA 1614
Db 1501 CCGAAGCCAGACTCACACAGCTGA 1524
RESULT 10
US-10-876-281-5
; Sequence 5, Application US/10876281
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-876-281-5
Query Match 89.6%; Score 1446.4; DB 22; Length 1840;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 1; Indels 33; Gaps 1;
Qy 124 GCCGCGCCCCCGCGCGGATGCTCTGTGCTGTATCTGAGATCATCGCGTTGAG 183
Db 1 GCCGCGCCCCCGCGCGGATGCTCTGTGCTGTATCTGAGATCATCGCGTTGAG 60
Qy 184 GAAACAGACGTTACGGGAAACATCAAGGCAAGTGGAAAAATGGCAAAATGGAAAGCCT 243
Db 61 GAAACAGACGTTACGGGAAACATCAAGGCAAGTGGAAAAATGGCAAAATGGAAAGCCT 120
Qy 244 TACGCTTTTACAGTTCACTGTGTAAAGAGACGACGACCGCTGGAAGTGGCGCAG 303
Db 121 TACGCTTTTACAGTTCACTGTGTAAAGAGACGACGACCGCTGGAAGTGGCGCAG 180
Qy 304 GTGACTTTTCTGGTGTCCAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCTCGGGAG 363
Db 181 GTGACTTTTCTGGTGTCCAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCTCGGGAG 240
Qy 364 ATGCTGGAAGAGCTGAGCTCCAGACCAAGCAATTTACTGGTATTATCAACCCGTTTGA 423
Db 241 ATGCTGGAAGAGCTGAGCTCCAGACCAAGCAATTTACTGGTATTATCAACCCGTTTGA 300
Qy 424 GGAAGAGCAAGCAAGCGGATATATGAAAGAAAAGTGGCACCACCTGTTCACTTAGCC 483
Db 301 GGAAGAGCAAGCAAGCGGATATATGAAAGAAAAGTGGCACCACCTGTTCACTTAGCC 360
Qy 484 TCCATCACCCTGACATCATCG-----TTACT 510
Db 361 TCCATCACCCTGACATCATCGGTAAACAAATTCATATGTTAACTATGTAGAAGTAATTA 420
Qy 511 GAAATGCTAATCAGGCCAAGAGACTCTGTATGAGATTACATAGACAAATACGACGGC 570
Db 421 GAAATGCTAATCAGGCCAAGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGC 480
Qy 571 ATCGTCTGTGCGGGAGATGGTATGTTTCAGCGAGGTGTCACGGGTCTGATTGGGAGG 630
Db 481 ATCGTCTGTGCGGGAGATGGTATGTTTCAGCGAGGTGTCACGGGTCTGATTGGGAGG 540

```
QY 631 ACGCAGAGAGCGCGGGTGCACAGAACACCCCGGGCTGTGTGTCCTCCAGTAGC 690
Db 541 ACGCAGAGAGCGCGGGTGCACAGAACACCCCGGGCTGTGTGTCCTCCAGTAGC 600
QY 691 CTCGGATTGGAATCATTTCCCGCAGGGTCAACGAGTCTGCTGTGTTACTTCCACCGTGGGC 750
Db 601 CTCGGATTGGAATCATTTCCCGCAGGGTCAACGAGTCTGCTGTGTTACTTCCACCGTGGGC 660
QY 751 ACCAGGAGCGAGAAACCTCGGGCTGCATATCTGTTGTTGGGACTCGTGGCCATGAT 810
Db 661 ACCAGGAGCGAGAAACCTCGGGCTGCATATCTGTTGTTGGGACTCGTGGCCATGAT 720
QY 811 GTGTCTCTAGTCCACACACAGCACACTCTCTGCTACTCCGTCCTCGTGGGCTAC 870
Db 721 GTGTCTCTAGTCCACACAGCACACTCTCTGCTACTCCGTCCTCGTGGGCTAC 780
QY 871 GGTCTTCTAGGGGACATCATCAAGCACAGTGAAGAAACGGTGGTCTTGGGCTTCCACAGA 930
Db 781 GGTCTTCTAGGGGACATCATCAAGCACAGTGAAGAAACGGTGGTCTTGGGCTTCCACAGA 840
QY 931 TAGGACTTTTCAAGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCC 990
Db 841 TAGGACTTTTCAAGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCC 900
QY 991 TTCTCTCTGCAACACACAGTGGGATCTCCAGGGATAGGAAGCCCTGCCGGGACAGA 1050
Db 901 TTCTCTCTGCAACACACAGTGGGATCTCCAGGGATAGGAAGCCCTGCCGGGACAGA 960
QY 1051 TGCTTTGTTTGCAGGCAAGCAGCAGCTGGAGGAGGAGCAAGAAAGCACTGTAT 1110
Db 961 TGCTTTGTTTGCAGGCAAGCAGCAGCTGGAGGAGGAGCAAGAAAGCACTGTAT 1020
QY 1111 GGTTTGGAAGCTGCGGAGGACGTGAGGAGTGGCAAGTCTGTGTGGAGTTTCTGGCC 1170
Db 1021 GGTTTGGAAGCTGCGGAGGACGTGAGGAGTGGCAAGTCTGTGTGGAGTTTCTGGCC 1080
QY 1171 ATCAATGCCACAAAATGCTGCTGTGTCGCGGAGCCCGAGGGGCTCTCCCGGCT 1230
Db 1081 ATCAATGCCACAAAATGCTGCTGTGTCGCGGAGCCCGAGGGGCTCTCCCGGCT 1140
QY 1231 GCCCACTTGGGAGAGCGGGTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTCAAT 1290
Db 1141 GCCCACTTGGGAGAGCGGGTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTCAAT 1200
QY 1291 TTTCTGAGATTCTCATCAGGACACCAACAGCAGGACAGTTTGAATTTCACTTTTGT 1350
Db 1201 TTTCTGAGATTCTCATCAGGACACCAACAGCAGGACAGTTTGAATTTCACTTTTGT 1260
QY 1351 GAAAGTTTATCGGTCAAGAAATTCAGTTTACGTGGAAGCAGTGGAGTGAAGACAGC 1410
Db 1261 GAAAGTTTATCGGTCAAGAAATTCAGTTTACGTGGAAGCAGTGGAGTGAAGACAGC 1320
QY 1411 GACCTCAAGAGGGGGGAGAGCGCTTTTGGGCAATTTGAGCAGACACCCCTCTGCG 1470
Db 1321 GACCTCAAGAGGGGGGAGAGCGCTTTTGGGCAATTTGAGCAGACACCCCTCTGCG 1380
QY 1471 TGCTGCACCGTCTCAACAGCTCTTGAATCTGGAACGCGGAGGTCCTGACAGCCCTGCC 1530
Db 1381 TGCTGCACCGTCTCAACAGCTCTTGAATCTGGAACGCGGAGGTCCTGACAGCCCTGCC 1440
QY 1531 ATCGAGGTCAGAGTCCATGCCAGCTGTTGCACTCTTTGCAAGGAAATGAAGAAAT 1590
Db 1441 ATCGAGGTCAGAGTCCATGCCAGCTGTTGCACTCTTTGCAAGGAAATGAAGAAAT 1500
QY 1591 CGGAAGCCAGACTCACACAGCTGA 1614
Db 1501 CGGAAGCCAGACTCACACAGCTGA 1524
```

RESULT 11

US-09-969-896-1

; Sequence 1, Application US/09969896

; Publication No. US20030125533A1

```
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: Kinase-Like Protein
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-1
```

Query Match 55.9%; Score 903; DB 10; Length 979;

Best Local Similarity 96.6%; Pred. No. 1.7e-268;

Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;

```
QY 387 ACCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGCAAGCAAGCGGAT 446
Db 1 ACCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGCAAGCAAGCGGAT 60
QY 447 ATATGAAAGAAAGTGGCACCACCTGTTCACTTACCTTACCTCCATCACCTGACATCATCG- 505
Db 61 ATATGAAAGAAAGTGGCACCACCTGTTCACTTACCTTACCTCCATCACCTGACATCATCG 120
QY 506 -----TTACTGAAACATGCTATCAGGCCAAGA 533
Db 121 TAACAAATTCATGTTAACTATGTAGAAAGTAAATTAATGAAATGCTAAATCAGGCCAAGA 180
QY 534 GACTCTGTATCAGATTAAACATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGG 593
Db 181 GACTCTGTATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGG 240
QY 594 TATGTTACGAGGCTGTGCACGGTCTGATTTGGGAGGACGACAGAGGAGCGCGGGGTGCA 653
Db 241 TATGTTACGAGGCTGTGCACGGTCTGATTTGGGAGGACGACAGAGGAGCGCGGGGTGCA 300
QY 654 CCAGAACACCCCGGGCTGTGCTGCCAGTAGCCTCCGGATTGGAATCATTTCCGCG 713
Db 301 CCAGAACACCCCGGGCTGTGCTGCCAGTAGCCTCCGGATTGGAATCATTTCCGCG 360
QY 714 AGGGTCAACGACGTCGCTGTACTCCACCGTGGGACACGACGACGAGAAACCTCGCG 773
Db 361 AGGGTCAACGACGTCGCTGTACTCCACCGTGGGACACGACGACGAGAAACCTCGCG 420
QY 774 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCTCAGTCCACCAACAG 833
Db 421 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCTCAGTCCACCAACAG 480
QY 834 CACACTCTTCGCTACTCCGCTGTCTGCTGGGCTTACGGCTTCTACGGGAGCATCATCA 893
Db 481 CACACTCTTCGCTACTCCGCTGTCTGCTGGGCTTCTACGGGAGCATCATCA 540
QY 894 GGACAGTGAAGAAACCGTGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 953
Db 541 GGACAGTGAAGAAACCGTGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 600
QY 954 CTTCTCTCCCAACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACACG 1013
Db 601 CTTCTCTCCCAACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACACG 660
QY 1014 GGGATCTCCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGTCAGGCAAAAGCA 1073
Db 661 GGGATCTCCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGTCAGGCAAAAGCA 720
QY 1074 GCAGCAGCTGGAGGAGGACGAGAGACACTGTATGTTTGGAGCTGGGAGGAGCT 1133
Db 1133
```

Db 721 GCAGCAGCTGGAGGAGGACAGAAAGACTGTATGTTTGAAGCTCGGAGGACGT 780
Qy 1134 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTCTG 1193
Db 781 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTCTG 840
Qy 1194 TGCTTGTGCGCGAGCCCCAGGGGCTCTCCCCGGCTGCCACTTGGGAGACGGGTCTTC 1253
Db 841 TGCTTGTGCGCGAGCCCCAGGGGCTCTCCCCGGCTGCCACTTGGGAGACGGGTCTTC 900
Qy 1254 TGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 1313
Db 901 TGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 960
Qy 1314 CACCAACGAGCAGGACGAG 1332
Db 961 CACCAACGAGCAGGACGAG 979

RESULT 12
US-10-631-958-1
; Sequence 1, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-1

Query Match 55.9%; Score 903; DB 19; Length 979;
Best Local Similarity 96.6%; Pred. No. 1.7e-268;
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;
Qy 387 ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACAAGGCAAGCGGAT 446
Db 1 ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACAAGGCAAGCGGAT 60
Qy 447 ATATGAAGAAAAGTGGCACCACCTGTTCACTTAGCTTCCATCACCACCTGACATCATCG- 505
Db 61 ATATGAAGAAAAGTGGCACCACCTGTTCACTTAGCTTCCATCACCACCTGACATCATCGG 120
Qy 506 -----TTACTGAACATGCTAATCAGGCCAAGGA 533
Db 121 TAACAAATTCATGTTAACTATGTAGAGTAATTAAGTCAATGCTAATCAGGCCAAGGA 180
Qy 534 GACTCTGTATGAGATTAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGG 593
Db 181 GACTCTGTATGAGATTAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGG 240
Qy 594 TATGTTACGAGGTGCTGCAAGGCTGATTTGGAGGACGAGAGGACGCGCGGGGTGCA 653
Db 241 TATGTTACGAGGTGCTGCAAGGCTGATTTGGAGGACGAGAGGACGCGCGGGGTGCA 300
Qy 654 CCAGAACCAACCCCGGGCTGTGCTGCCAGTAGCTCCGGATTGGAATCATTTCCCGC 713
Db 301 CCAGAACCAACCCCGGGCTGTGCTGCCAGTAGCTCCGGATTGGAATCATTTCCCGC 360
Qy 714 AGGGTCAACGGAATGCTGCTGTTACTCCACCGTGGGCCACGACGAGCAACCTCGGC 773

Db 361 AGGGTCAACGGAATGCTGCTGTACTCCAGCGTGGCACCAGCGACGAGAAACCTCGGC 420
Qy 774 GCTGCATATCGTTGTTGGGACCTCGCTGGCCATGATGTCTCCTCAGTCCACCAACAACAG 833
Db 421 GCTGCATATCGTTGTTGGGACCTCGCTGGCCATGATGTCTCCTCAGTCCACCAACAACAG 480
Qy 834 CACACTCTCTCGTACTCCGCTGCTCGCTGGGCTACGGCTTCTACGGGACATCATCAA 893
Db 481 CACACTCTCTCGTACTCCGCTGCTCGCTGGGCTACGGCTTCTACGGGACATCATCAA 540
Qy 894 GGACAGTGAAGAAAACGGTGGTTCCTGTCAGATACGACTTTTTCAGGTTTAAAGAC 953
Db 541 GGACAGTGAAGAAAACGGTGGTTCCTGTCAGATACGACTTTTTCAGGTTTAAAGAC 600
Qy 954 CTTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTCCTCCCTGCACAACACACGGT 1013
Db 601 CTTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTCTCCTGCACAACACACGGT 660
Qy 1014 GGGATCTCCAAGGATAGGAAGCCCTCGCGGACAGGATGCTTTTTCGAGGCAAAAGCAA 1073
Db 661 GGGATCTCCAAGGATAGGAAGCCCTCGCGGACAGGATGCTTTTTCGAGGCAAAAGCAA 720
Qy 1074 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGT 1133
Db 721 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGT 780
Qy 1134 GGAGGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTCTG 1193
Db 781 GGAGGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTCTG 840
Qy 1194 TGCTTGTGCGCGAGCCCCAGGGGCTCTCCCCGGCTGCCACACTTGGGAGACGGGTCTTC 1253
Db 841 TGCTTGTGCGCGAGCCCCAGGGGCTCTCCCCGGCTGCCACACTTGGGAGACGGGTCTTC 900
Qy 1254 TGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1313
Db 901 TGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTTCTCATCAGGCA 960
Qy 1314 CACCAACGAGCAGGACGAG 1332
Db 961 CACCAACGAGCAGGACGAG 979

RESULT 13
US-10-635-247
; Sequence 247, Application US/10115635
; Publication No. US20040137434A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 797CON
; CURRENT APPLICATION NUMBER: US/10/115.635
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 362
; SOFTWARE: PE_FL_genes Version 2.0
; SEQ ID NO 247
; LENGTH: 817
; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (90)..(662)
US-10-115-635-247

Query Match 35.3%; Score 569; DB 19; Length 817;

Best Local Similarity 100.0%; Pred. No. 4.4e-165; Mismatches 0; Indels 0; Gaps 0;

Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGCGACGGGGGCGGAGCGCGTGCATCCGTGCTGTGGTGAAGCAGCAGCGC 60

DB 90 ATGGGGGCGACGGGGGCGGAGCGCGTGCATCCGTGCTGTGGTGAAGCAGCAGCGC 149

QY 61 TCGCCGCTGAGCCTGGAGCCCGCGCGCTCTGCTGCGCTGTGCGCGAGCCCGGGGCC 120

DB 150 TCGCCGCTGAGCCTGGAGCCCGCGCGCTCTGCTGCGCTGTGCGCGAGCCCGGGGCC 209

QY 121 GGAGCGGGCGCCCGCGCGGATCGCTCTGTGCTGTGCTGTATCTGAGATCATCGCGTT 180

DB 210 GGAGCGGGCGCCCGCGCGGATCGCTCTGTGCTGTGCTGTATCTGAGATCATCGCGTT 269

QY 181 GAGGAAACAGAGCTTCAAGGAAACATCAAGGCACTGGAAATGCGAAATGGAAG 240

DB 270 GAGGAAACAGAGCTTCAAGGAAACATCAAGGCACTGGAAATGCGAAATGGAAG 329

QY 241 CCTTACGCTTTTACAGTTCACTGTGTAAGAGAGCAGCAGCGCCGCTGGAAGTGGCG 300

DB 330 CCTTACGCTTTTACAGTTCACTGTGTAAGAGAGCAGCAGCGCCGCTGGAAGTGGCG 389

QY 301 CAGGTGACTTTTGTGTGTCAGAGAGCAGCTGTCTACTTGTGCTGAGACCCCTGGCG 360

DB 390 CAGGTGACTTTTGTGTGTCAGAGAGCAGCTGTCTACTTGTGCTGAGACCCCTGGCG 449

QY 361 GAGTGTGGAGAGCTGACGTCAGACCAAGCAATTTACTGTTATTTATCAACCCGTTT 420

DB 450 GAGTGTGGAGAGCTGACGTCAGACCAAGCAATTTACTGTTATTTATCAACCCGTTT 509

QY 421 GGAGAAAGGCAAGGCAAGCGGATATATGAAGAAAGTGGCACTGTTCACCTTA 480

DB 510 GGAGAAAGGCAAGGCAAGCGGATATATGAAGAAAGTGGCACTGTTCACCTTA 569

QY 481 GCCTCATCACCACATGACATCATGTTACTGAACATGCTAATCAGGCCAAGAGACTCTG 540

DB 570 GCCTCATCACCACATGACATCATGTTACTGAACATGCTAATCAGGCCAAGAGACTCTG 629

QY 541 TATGAGATTAAACATAGACAAATACGACG 569

DB 630 TATGAGATTAAACATAGACAAATACGACG 658

RESULT 14

US-09-969-896-8

; Sequence 8, Application US/09969896

; Publication No. US20030125533A1

; GENERAL INFORMATION:

; APPLICANT: Kossida, Sophia

; TITLE OF INVENTION: Regulation of human Sphingosine

; FILE REFERENCE: 004974.00594

; CURRENT APPLICATION NUMBER: US/09/969,896

; PRIOR FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: US 60/238,005

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: US 60/314,113

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 550

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-969-896-8

Query Match 29.4%; Score 475; DB 10; Length 550;

Best Local Similarity 100.0%; Pred. No. 4.8e-136;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGCGACGGGGGCGGAGCGCGTGCATCCGTGCTGTGGTGAAGCAGCAGCGC 60

DB 76 ATGGGGGCGACGGGGGCGGAGCGCGTGCATCCGTGCTGTGGTGAAGCAGCAGCGC 135

QY 61 TGCSCCTGAGCCTGGAGCCCGCGGCTCTGCTGCGCTGTGCGAGCCCGGGGCC 120

DB 136 TGCSCCTGAGCCTGGAGCCCGCGGCTCTGCTGCGCTGTGCGAGCCCGGGGCC 195

QY 121 GGAGCGGGCGCCCGCGCGGATCGCTCTGTGCTGTATCTGAGATCATCGCGTT 180

DB 196 GGAGCGGGCGCCCGCGCGGATCGCTCTGTGCTGTATCTGAGATCATCGCGTT 255

QY 181 GAGGAAACAGAGCTTCAAGGAAACATCAAGGCACTGGAAATGCGAAATGGAAG 240

DB 256 GAGGAAACAGAGCTTCAAGGAAACATCAAGGCACTGGAAATGCGAAATGGAAG 315

QY 241 CCTTACGCTTTTACAGTTCACTGTGTAAGAGAGCAGCAGCGCCGCTGGAAGTGGCG 300

DB 316 CCTTACGCTTTTACAGTTCACTGTGTAAGAGAGCAGCAGCGCCGCTGGAAGTGGCG 375

QY 301 CAGGTGACTTTTGTGTGTCAGAGAGCAGCTGTCTACTTGTGCTGAGACCCCTGGCG 360

DB 376 CAGGTGACTTTTGTGTGTCAGAGAGCAGCTGTCTACTTGTGCTGAGACCCCTGGCG 435

QY 361 GAGTGTGGAGAGCTGACGTCAGACCAAGCAATTTACTGTTATTTATCAACCCGTTT 420

DB 436 GAGTGTGGAGAGCTGACGTCAGACCAAGCAATTTACTGTTATTTATCAACCCGTTT 495

QY 421 GGAGAAAGGCAAGGCAAGCGGATATATGAAGAAAGTGGCACTGTTC 475

DB 496 GGAGAAAGGCAAGGCAAGCGGATATATGAAGAAAGTGGCACTGTTC 550

RESULT 15

US-10-631-958-8

; Sequence 8, Application US/10631958

; Publication No. US20040192580A1

; GENERAL INFORMATION:

; APPLICANT: Kossida, Sophia

; TITLE OF INVENTION: Regulation of human Sphingosine

; FILE REFERENCE: 004974.00594

; CURRENT APPLICATION NUMBER: US/10/631,958

; CURRENT FILING DATE: 2003-08-01

; PRIOR APPLICATION NUMBER: US/09/969,896

; PRIOR FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: US 60/238,005

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: US 60/314,113

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 550

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-631-958-8

Query Match 29.4%; Score 475; DB 19; Length 550;

Best Local Similarity 100.0%; Pred. No. 4.8e-136;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGCGACGGGGGCGGAGCGCGTGCATCCGTGCTGTGGTGAAGCAGCAGCGC 60

DB 76 ATGGGGGCGACGGGGGCGGAGCGCGTGCATCCGTGCTGTGGTGAAGCAGCAGCGC 135

QY 61 TGCSCCTGAGCCTGGAGCCCGCGGCTCTGCTGCGCTGTGCGAGCCCGGGGCC 120

Db	136	TGCGCCGTGAGCCCTGGAGCCCGCGGGCTCTGCTGCGCTGGTGGCGGAGCCCGGGGCC	195
Qy	121	GGAGCCGGCGCCCGCGGGATGCTGTGCTGTGCTGTATCTGAGATCATCGCGTT	180
Db	196	GGAGCCGGCGCCCGCGGGATGCTGTGCTGTGCTGTATCTGAGATCATCGCGTT	255
Qy	181	GAGGAAACAGACGTTTCACGGGAAACATCAAGGCAGTGGAAAATGGCAGAAAATGAAAAAG	240
Db	256	GAGGAAACAGACGTTTCACGGGAAACATCAAGGCAGTGGAAAATGGCAGAAAATGAAAAAG	315
Qy	241	CTTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGGCACCGTGGAAAGTGGCG	300
Db	316	CTTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGGCACCGTGGAAAGTGGCG	375
Qy	301	CAGGTGACTTTCTGCTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCTTGGG	360
Db	376	CAGGTGACTTTCTGCTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCTTGGG	435
Qy	361	GAGATGCTGGAGAAAGCTGACCTCCAGACCAAGCATTCTGTTATTTATCAACCCGTTT	420
Db	436	GAGATGCTGGAGAAAGCTGACCTCCAGACCAAGCATTCTGTTATTTATCAACCCGTTT	495
Qy	421	GGAGGAAAGGACAAAGCGGATATATGAAAGAAAAGTGGCACCACTGTTCA	475
Db	496	GGAGGAAAGGACAAAGCGGATATATGAAAGAAAAGTGGCACCACTGTTCA	550

Search completed: September 5, 2005, 20:00:16
Job time : 1181.31 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 03:56:01 ; Search time 17.8442 Seconds
(without alignments)
1363.781 Million cell updates/sec

Title: , US-10-631-958-2
Perfect score: 1717
Sequence: 1 PKHLVFINPGCKGQKRI.....KCSRNFURFLRHTNQDQ 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCITUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	613	35.7	687	4	US-09-270-767-45874
2	566	33.0	359	4	US-09-270-767-46720
3	282	16.4	490	4	US-10-053-510-19
4	270.5	15.8	524	4	US-10-053-510-20
5	243	14.2	618	4	US-09-970-516-4
6	243	14.2	618	4	US-09-817-676A-14
7	236	13.7	617	4	US-09-817-676A-12
8	228	13.3	384	4	US-09-949-016-7026
9	228	13.3	384	4	US-09-796-487-3
10	227.5	13.2	384	4	US-09-970-516-2
11	227.5	13.2	384	4	US-09-959-897-2
12	226	13.2	373	4	US-09-796-487-5
13	226	13.2	381	4	US-09-796-487-1
14	226	13.2	381	4	US-09-796-487-4
15	226	13.2	382	4	US-09-970-516-6
16	226	13.2	388	4	US-09-817-676A-15
17	226	13.2	388	4	US-09-796-487-2
18	223.5	13.0	368	4	US-10-053-510-21
19	220.5	12.8	424	4	US-09-796-487-8
20	219.5	12.8	392	4	US-09-796-487-6
21	216	12.6	536	4	US-09-248-796A-15859
22	193	11.2	204	4	US-09-796-487-9
23	181.5	10.6	403	4	US-09-796-487-7
24	172	10.0	299	4	US-09-270-767-61405
25	143	8.3	312	4	US-09-949-016-9811
26	132	7.7	313	4	US-09-107-532A-7154
27	131.5	7.7	372	4	US-09-603-208A-264

28	129.5	7.5	119	4	US-09-205-258-788	Sequence 788, App
29	125	7.3	294	4	US-09-583-110-3740	Sequence 3740, Ap
30	125	7.3	315	4	US-09-107-433-3121	Sequence 3121, Ap
31	118.5	6.9	283	4	US-09-634-238-373	Sequence 373, App
32	117	6.8	160	4	US-09-270-767-62339	Sequence 62339, A
33	117	6.8	304	4	US-09-134-000C-5366	Sequence 5366, Ap
34	110	6.4	303	4	US-09-107-532A-5133	Sequence 5133, Ap
35	109	6.3	325	4	US-09-107-532A-6786	Sequence 6786, Ap
36	108	6.3	293	4	US-09-205-258-328	Sequence 328, App
37	106	6.2	316	4	US-09-710-279-1350	Sequence 1350, Ap
38	106	6.2	324	3	US-09-134-001C-4774	Sequence 4774, Ap
39	105	6.1	345	4	US-09-134-000C-3813	Sequence 3813, Ap
40	104	6.1	63	4	US-09-959-897-45	Sequence 45, Appl
41	102	5.9	64	4	US-09-959-897-37	Sequence 37, Appl
42	92	5.4	65	4	US-09-959-897-53	Sequence 53, Appl
43	90.5	5.3	64	4	US-09-959-897-29	Sequence 29, Appl
44	89.5	5.2	323	4	US-09-107-532A-5598	Sequence 5598, Ap
45	89	5.2	297	4	US-09-134-000C-3616	Sequence 3616, Ap

ALIGNMENTS

RESULT 1

US-09-270-767-45874
; Sequence 45874, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45874
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-45874

Query Match	35.7%	Score 613;	DB 4;	Length 687;
Best Local Similarity	35.2%	Pred. No. 1.3e-59;		
Matches 134;	Conservative 59;	Mismatches 108;	Indels 80;	Gaps 7;
QY	2	KHLVFINPGCKGQKRIYERKVAFLFTLASITTTDIIGNKFYVNYVEVITEHANQAKET	61	
Db	204	RLLVFINPYGRKAGATYERHVRPIQLAGVATCI-----TTORANQVKDI	252	
QY	62	LYEINIDKYDGVCGDGMFSEVLHGLIGRTQTSAGVDQNHPRAVLVPSSLRIGIIPAG	121	
Db	253	LLSHDLGVYDVCVCGDGTVAEVINGLIFRQRELGLDEQRPPIPRP-ALPVGVIIPAG	311	
QY	122	STDCVCYSTGTSDAETSAHLIVVGDLSLAMDVSSVHHNSTLLRYSVSLGCVGYGDIIDK	181	
Db	312	STDTIAYSMHGTADVTAAIHVILGQHRGLDVCVSNQGLLRFCASVLSYGLDVAQA	371	
QY	182	SEKKRWLGLARYDFSGLTKFLSHHCYECTVSFL-----PAQHTVGSPRD-----	225	
Db	372	SENRYWMPRYEYSGVKAFLLNNGYDAELRWLKEPDLTLTTPLEDIFQSPDVSCLGES	431	
QY	226	-RKPCRACGFCV-----ROSKQOQLEEEOKKALYGLAEADVE-----	261	
Db	432	VPSVCYANCQRCSPASSIQEQRSSLIQESKEA----ERNQOVETEDSHLAASEAALLR	487	
QY	262	-----EMQVVCCKFLAINATNMSCACRRSPRGLSPAHLGD	297	
Db	488	PRPRPGNRLRPTGSISSMRNLGNQDWKYVRGNFFMICGATTCACARSPNGLSRYSHLGD	547	
QY	298	GSSDLILIRKCSRNFRLFLI	318	
Db	548	GCLDLILVKKTSLNNVRPLL	568	


```
QY 226 RKPCRACF-----VCRSQKQLEEE-----OK 248
Db 254 RMQSRSCNTHDMLNGPAPIYHSSAEYLPQEFADVLSLETINSQSRCDSWLGGSR 313
QY 249 KALY-----GLEAAE-----DVEE-----WQVWCG 268
Db 314 RSFYYSISBSIYHSLADESEFAGLAAASLENQQNYGPASELPDLNEPSEBQGLVREG 373
QY 269 KFLAINA-----TNMSACRRSPRGLGPAHGLDGGSDLLIRK-CRSENFRL 317
Db 374 EFWMHVVYQTHGIDCH-----FAPKAQLNDGTIYLILIRAGISRPHELLSEL 421

RESULT 5
US-09-970-516-4
; Sequence 4, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: polynucleotides encoding sphingosine kinases
; CURRENT APPLICATION NUMBER: 4-31617
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-516-4

Query Match 14.2%; Score 243; DB 4; Length 618;
Best Local Similarity 28.9%; Pred. No. 4.6e-18;
Matches 67; Conservative 43; Mismatches 90; Indels 32; Gaps 4;

QY 1 PKHLVFNPFQGGKQKRIYERKVPAPLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
Db 144 PPRLLLVNPFQGGKQKRIYERKVPAPLFTLASITTDIIGNKFYVYVEVITEHANOAKE 192
QY 61 TLYENIDKDGIVCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120
Db 193 LVQGLSLEWDGIVTVSGDGLLHEVLGLLDR-----PDWEAVKMP-----VGIILPC 240
QY 121 GSTDCVCYS-----TVGTSDAETSALHIWVGDSLAMDVSSVHHNSTLLRYSVSLIG 171
Db 241 GSGNALAGAVNQHGGEFEPALGLDLLNCSLLLCRGGHPDLLSVTLASGSRCSFSLVA 300
QY 172 YGFYGDIIKDEKKRWGLGARYDFSGLTFLSHHCYEGTVSFLPAQHTVGP 223
Db 301 WGFVSDVDIQSERFRALGSARFTLTGLVLGLATLHTYRGRLSYLPATVPEPASP 352

RESULT 6
US-09-817-676A-14
; Sequence 14, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: Expression and Methods of Use Thereof
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-817-676A-14

Query Match 13.7%; Score 236; DB 4; Length 617;
Best Local Similarity 29.7%; Pred. No. 2.8e-17;
Matches 66; Conservative 40; Mismatches 84; Indels 32; Gaps 4;

QY 4 LLVFINPFQGGKQKRIYERKVPAPLFTLASITTDIIGNKFYVYVEVITEHANOAKETLY 63
Db 148 LLILVNPFGGRGLAWQRCMDHVPWPMISEAGLSFNLIQ-----TERQNHARELVQ 196
QY 64 EINTDKYDGIYCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGST 123
Db 197 GLSLSEWEGIVTVSGDGLLYEVLNGLLDR-----PDWEDAVKMP-----IGVLPCGSG 244
QY 124 DVCV-----YSTVGTSDAETSALHIWVGDSLAMDVSSVHHNSTLLRYSVSLIGYGF 174
Db 245 NALAGAVSHHGGFEQVGVVDLLNCSLLLCRGGHPDLLSVTLASGSRCSFSLVANGF 304
QY 175 YGDIKDEKKRWGLGARYDFSGLTFLSHHCYEGTVSFLPA 216
Db 305 LSDVDIHSERFRALGSARFTLTGLVGLASLTHTYRGRLSYLPA 346

RESULT 8
US-09-949-016-7026
; Sequence 7026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
QY 226 RKPCRACF-----VCRSQKQLEEE-----OK 248
Db 254 RMQSRSCNTHDMLNGPAPIYHSSAEYLPQEFADVLSLETINSQSRCDSWLGGSR 313
QY 249 KALY-----GLEAAE-----DVEE-----WQVWCG 268
Db 314 RSFYYSISBSIYHSLADESEFAGLAAASLENQQNYGPASELPDLNEPSEBQGLVREG 373
QY 269 KFLAINA-----TNMSACRRSPRGLGPAHGLDGGSDLLIRK-CRSENFRL 317
Db 374 EFWMHVVYQTHGIDCH-----FAPKAQLNDGTIYLILIRAGISRPHELLSEL 421

RESULT 5
US-09-970-516-4
; Sequence 4, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: polynucleotides encoding sphingosine kinases
; CURRENT APPLICATION NUMBER: 4-31617
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-516-4

Query Match 14.2%; Score 243; DB 4; Length 618;
Best Local Similarity 28.9%; Pred. No. 4.6e-18;
Matches 67; Conservative 43; Mismatches 90; Indels 32; Gaps 4;

QY 1 PKHLVFNPFQGGKQKRIYERKVPAPLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
Db 144 PPRLLLVNPFQGGKQKRIYERKVPAPLFTLASITTDIIGNKFYVYVEVITEHANOAKE 192
QY 61 TLYENIDKDGIVCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120
Db 193 LVQGLSLEWDGIVTVSGDGLLHEVLGLLDR-----PDWEAVKMP-----VGIILPC 240
QY 121 GSTDCVCYS-----TVGTSDAETSALHIWVGDSLAMDVSSVHHNSTLLRYSVSLIG 171
Db 241 GSGNALAGAVNQHGGEFEPALGLDLLNCSLLLCRGGHPDLLSVTLASGSRCSFSLVA 300
QY 172 YGFYGDIIKDEKKRWGLGARYDFSGLTFLSHHCYEGTVSFLPAQHTVGP 223
Db 301 WGFVSDVDIQSERFRALGSARFTLTGLVLGLATLHTYRGRLSYLPATVPEPASP 352

RESULT 6
US-09-817-676A-14
; Sequence 14, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: Expression and Methods of Use Thereof
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 618
; TYPE: PRT
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7026
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7026

Query Match      13.3%; Score 228; DB 4; Length 384;
Best Local Similarity 25.9%; Pred. No. 1e-16;
Matches 88; Conservative 53; Mismatches 137; Indels 62; Gaps 13;

Qy 1 PKHLLVFNPFGGKGQKRIYERKVPAPLFTLASITTDIIGNKFVYVYVEVITEHANOAKE 60
Db 14 PCRVLLNPRGGKGKALQFRSHVQPLLAEEI-----SFTLMLTERRNHARE 62

Qy 61 TLYEINDKYDGVGVGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPA 120
Db 63 LVRSEELGRDALVVMGSDGLMHEVYVNGLMERPDPWETAIQK-----PLCSLIPA 110

Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLLG 171
Db 111 GSGNALAASLNHYAGYEQVTNEDLLTCTLLCRLLSPMULLSLHTASGLRFLSVLSLA 170

Qy 172 YGFYGDIIKDEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRA 231
Db 171 WGFIADVLESEKRYRLGEMRFTLTGTLRLAALRTYRGLAYLPVGR-VGSKTPASP--- 226

Qy 232 GCFVCRQSKQOLEEQQKALVGLEAAEDV-EWQVVCVK-----FLAINATMNSCACRRSP 286
Db 227 --VVVQOGPVDAAH-----LVPLE--EPVPSHTWMPDDEDFVLILALHSLGSEMFAAP 276

Qy 287 RGLSPAHLGGSSDLILIRK-CSRFNFRFLI-----RH 320
Db 277 MRCRA-----GVMLFVVRAGVSRAMLLRFLAMEKGRH 311

RESULT 9
US-09-796-487-3
; Sequence 3, Application US/09796487
; Patent No. 6830916
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(384)
; OTHER INFORMATION: SEQ ID NO 3 is the peptide sequence of hSPHK1 in Fig. 3, correspondi
; OTHER INFORMATION: nding to amino acid residue 1 to 384 of Homo sapiens SPHK-1 of Ge
```

```
; OTHER INFORMATION: nBank sequence Accession Number AAF73423.
; PUBLICATION INFORMATION:
; AUTHORS: Nava et al.
; TITLE: Functional characterization of human sphingosine kinase-1
; JOURNAL: FEBS Lett.
; VOLUME: 473
; ISSUE: 1
; PAGES: 81-84
; DATE: 2000
; DATABASE ACCESSION NUMBER: AAF73423
; DATABASE ENTRY DATE: 2000-06-01
; RELEVANT RESIDUES: (1)..(384)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAF73423
; DATABASE ENTRY DATE: 2000-06-01
; RELEVANT RESIDUES: (1)..(384)
US-09-796-487-3

Query Match      13.3%; Score 228; DB 4; Length 384;
Best Local Similarity 25.9%; Pred. No. 1e-16;
Matches 88; Conservative 53; Mismatches 137; Indels 62; Gaps 13;

Qy 1 PKHLLVFNPFGGKGQKRIYERKVPAPLFTLASITTDIIGNKFVYVYVEVITEHANOAKE 60
Db 14 PCRVLLNPRGGKGKALQFRSHVQPLLAEEI-----SFTLMLTERRNHARE 62

Qy 61 TLYEINDKYDGVGVGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPA 120
Db 63 LVRSEELGRDALVVMGSDGLMHEVYVNGLMERPDPWETAIQK-----PLCSLIPA 110

Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLLG 171
Db 111 GSGNALAASLNHYAGYEQVTNEDLLTCTLLCRLLSPMULLSLHTASGLRFLSVLSLA 170

Qy 172 YGFYGDIIKDEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRA 231
Db 171 WGFIADVLESEKRYRLGEMRFTLTGTLRLAALRTYRGLAYLPVGR-VGSKTPASP--- 226

Qy 232 GCFVCRQSKQOLEEQQKALVGLEAAEDV-EWQVVCVK-----FLAINATMNSCACRRSP 286
Db 227 --VVVQOGPVDAAH-----LVPLE--EPVPSHTWMPDDEDFVLILALHSLGSEMFAAP 276

Qy 287 RGLSPAHLGGSSDLILIRK-CSRFNFRFLI-----RH 320
Db 277 MRCRA-----GVMLFVVRAGVSRAMLLRFLAMEKGRH 311

RESULT 10
US-09-796-516-2
; Sequence 2, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-516-2

Query Match      13.2%; Score 227.5; DB 4; Length 384;
Best Local Similarity 27.0%; Pred. No. 1.2e-16;
Matches 64; Conservative 43; Mismatches 97; Indels 33; Gaps 5;

Qy 1 PKHLLVFNPFGGKGQKRIYERKVPAPLFTLASITTDIIGNKFVYVYVEVITEHANOAKE 60
Db 14 PCRVLLNPRGGKGKALQFRSHVQPLLAEEI-----SFTLMLTERRNHARE 62
```

Qy	61	TLYEINIDKYDGVTCVGGDGMSEVLHGLIGRTQBSAGVDQNHPRAVLVFSSIRIGIIPA	120
	:	:	:
Db	63	LVRESBLSGRWDALVWNSGDGLMHEVYVGLMERPDWETAIQK-----PLCSLPA	110
	:	:	:
Qy	121	GSTDVCV-----YSTVGTSDAETSAHLTVVGDLSA-MDVSSVHHNSTLLRKYSVSLIG	171
	:	:	:
Db	111	GSGNLAASLHHVAGYEQVNTNEDLLTNCITLLCRRLSPMNLISLHTVASSGRVFSVL	170
	:	:	:
Qy	172	YGYPGDIIKDSEKKRWGLIARYDFSLGKTFIISHHCYEGTVSFIPAQHTVGSPRDRXP	228
	:	:	:
Db	171	WGFIADVDLSESKYRRLGEMRFTLTGFTLRLAALRTYGRILAYLPVGR-VGSKTPASP	226
	:	:	:

RESULT 11

US-09-953-897-2
 ; Sequence 2, Application US/09959897
 ; Patent No. 6730480
 ; GENERAL INFORMATION:
 ; APPLICANT: PITSON, Stuart M
 ; APPLICANT: Brian, WATTENBERG W
 ; APPLICANT: Fu, Xia
 ; APPLICANT: Richard, D'ANDREA J
 ; APPLICANT: Jennifer, BAMBLE R
 ; APPLICANT: Mathew, VADAS A
 ; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
 ; FILE REFERENCE: PITSON=1
 ; CURRENT APPLICATION NUMBER: US/09/959,897
 ; CURRENT FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: PCT/AU00/00457
 ; PRIOR FILING DATE: 2000-05-12
 ; PRIOR APPLICATION NUMBER: AU PQ 0339
 ; PRIOR FILING DATE: 1999-05-13
 ; PRIOR APPLICATION NUMBER: AU PQ 1504
 ; PRIOR FILING DATE: 1999-07-08
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 384
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-959-897-2

Query Match	13.2%	Score 227.5;	DB 4;	Length 384;
Best Local Similarity	27.0%;	Pred. No. 1.2e-16;		
Matches	64;	Conservative 43;	Mismatches 97;	Indels 33; Gaps 5;

Qy	1	P	K	H	L	L	V	I	N	P	G	G	K	G	K	R	I	Y	E	R	V	K	V	A	P	L	T	T	L	A	S	I	T	T	D	I	I	G	N	K	F	Y	V	V	E	V	I	T	E	H	A	N	O	A	K	E	60
Db	14	P	C	R	V	L	L	N	P	R	G	G	K	A	L	O	L	F	R	S	H	V	O	P	L	A	E	I	-----	S	F	T	M	L	T	E	R	N	H	A	R	E	62														
Qy	61	T	L	E	I	N	D	I	K	D	G	I	V	C	G	D	M	F	S	E	L	L	H	L	I	G	T	O	R	S	A	G	V	Q	N	H	P	R	A	V	L	P	S	S	L	R	I	G	I	P	A	120					
Db	63	L	V	R	S	E	E	L	G	R	W	D	A	L	V	M	S	D	G	L	M	H	E	V	N	G	L	M	R	P	D	E	T	A	I	O	K	-----	P	L	C	S	L	P	A	110											
Qy	121	G	S	T	D	C	V	C	-----	Y	S	T	V	G	T	S	D	A	E	T	A	H	I	V	G	D	S	L	A	-	M	D	V	S	S	V	H	N	S	T	L	R	Y	S	V	S	L	L	G	171							
Db	111	G	S	G	N	A	L	A	S	N	H	Y	A	G	Y	E	O	V	T	N	E	D	L	L	T	N	C	T	L	C	R	L	L	S	P	M	L	L	S	H	T	A	G	S	L	R	F	S	V	L	S	L	A	170			
Qy	172	Y	G	F	G	D	I	I	K	O	S	K	K	R	W	L	G	A	R	Y	D	F	S	G	L	T	F	L	S	H	H	C	Y	E	G	T	V	S	F	L	P	A	O	H	V	T	G	S	P	R	K	P	228				
Db	171	W	G	F	T	A	D	V	D	L	S	E	K	V	R	L	G	E	M	R	F	T	G	T	F	L	R	A	A	L	T	Y	R	G	L	A	Y	L	P	V	G	R	-	V	G	S	K	T	P	A	S	P	226				

RESULT 12

```

US-09-796-487-5
; Sequence 5, Application US/09796487
; Patent No. 6830916
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Spingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796.487

```

```

; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: (1)_(373)
; LOCATION: (1)_(373)
; OTHER INFORMATION: SEQ ID NO 5 is the peptide sequence of Mouse SPK1a in F
; OTHER INFORMATION: responding to amino acid residue 131 to 504 of SPK1a o
; OTHER INFORMATION: sequence Accession Number AAC61697.
; PUBLIC INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (132)..(504)
US-09-796-487-5

Query Match 13.2%; Score 226; DB 4; Length 373;
Best Local Similarity 26.6%; Pred. No. 1.7e-16;
Matches 63; Conservative 44; Mismatches 94; Indels 36; Gaps 5;

Qy 1 PKHLLVFIPNPGKGQGRKRIYERKVAPLFTLASITTTDIGNKFYVNYVEVITEHANOAKE 60
Db 5 PCRVLVLLNPGQGGKALQIFQRVQPFLEEAITFKLI-----LTERKNHARE 53
Qy 61 TLYEINIDKYDGVICVCGGDMGFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120
Db 54 LVCAEELGHWDALAVMSGDGLMEVNVGLMERPDWETAIQK-----PLCSLPG 101
Qy 121 GSTDCVC-----YSTVCTSDAETSALHIVGDSLIA-MDVSVVHNSTLLRYSVSLLG 171
Db 102 GSGNALAASVNHVAGYEQVNTDGLLINCTLLCCRRLSPMNLISLHTAGSLRLYSLSLS 161
Qy 172 YGFVGDIIKDSEKKRWGLIARYDFSGLKTFLSHHCYEGVTVSFLPAQHTVGSPRDRKP 228
Db 162 WGFVADVDLSEKRYRIGRTFTVGTGTFPRISLIYOGOLAVLP-----VGTWASKRP 214

```

RESIST. T. 13

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```
; TITLE: Molecular cloning and functional characterization of murine sphingosin
; TITLE: kinase
; JOURNAL: Journal of Biological Chemistry
; VOLUME: 237
; ISSUE: 37
; PAGES: 23722-23782
; DATE: 1998
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (124)..(504)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (124)..(504)
US-09-796-487-1

Query Match          13.2%; Score 226; DB 4; Length 381;
Best Local Similarity 26.6%; Pred. No. 1.7e-16;
Matches 63; Conservative 44; Mismatches 94; Indels 36; Gaps 5;

Qy 1 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
Db 13 PCRVLLNPOGGKGKALQLQFSRVQPFLEAEITFKLI-----LTERKNHARE 61

Qy 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120
Db 62 LVCAEELGHWDAVMGSDGLMHVVGMLMERPDWETAIOK-----PLCSLPG 109

Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLIG 171
Db 110 GSGNALAASVNHVAGYEQVTNEDLLINCTLLCRRRLSPMNLLSLHTASGLRLYSVLSLS 169

Qy 172 YGFYGDIIKDKSEKRWLGARYDFSGLKTFLSHHCYEGTVSFFLPAQHTVGSPPDRKP 228
Db 170 WGFVADVLESEKRYRLGEIRFTVGTFFRLASLRIOQLAYLP-----VGTVASKRP 222

RESULT 14
US-09-796-487-4
; Sequence 4, Application US/09796487
; Patent No. 6830916
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(381)
; OTHER INFORMATION: SEQ ID NO 4 is the peptide sequence of mSPHK1a in Figure 3, corre
; OTHER INFORMATION: sponding to amino acid residue 124 to 504 of SPHK1a of GenBank se
; OTHER INFORMATION: quence Accession Number AF068748. SEQ ID NO 4 is equivalent to SE
; OTHER INFORMATION: Q ID NO 1 that is the amino acid sequence of SPHK1a in Figure 1.
; PUBLICATION INFORMATION:
; AUTHORS: Kohama et al.,
; TITLE: Molecular cloning and functional characterization of murine sphingosine
; JOURNAL: Journal of Biological Chemistry
; VOLUME: 273
; ISSUE: 37
; PAGES: 23722-23782
; DATE: 1998
```

```
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (124)..(504)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (124)..(504)
US-09-796-487-4

Query Match          13.2%; Score 226; DB 4; Length 381;
Best Local Similarity 26.6%; Pred. No. 1.7e-16;
Matches 63; Conservative 44; Mismatches 94; Indels 36; Gaps 5;

Qy 1 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
Db 13 PCRVLLNPOGGKGKALQLQFSRVQPFLEAEITFKLI-----LTERKNHARE 61

Qy 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120
Db 62 LVCAEELGHWDAVMGSDGLMHVVGMLMERPDWETAIOK-----PLCSLPG 109

Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLIG 171
Db 110 GSGNALAASVNHVAGYEQVTNEDLLINCTLLCRRRLSPMNLLSLHTASGLRLYSVLSLS 169

Qy 172 YGFYGDIIKDKSEKRWLGARYDFSGLKTFLSHHCYEGTVSFFLPAQHTVGSPPDRKP 228
Db 170 WGFVADVLESEKRYRLGEIRFTVGTFFRLASLRIOQLAYLP-----VGTVASKRP 222

RESULT 15
US-09-970-516-6
; Sequence 6, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: polynucleotides encoding sphingosine kinases
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-970-516-6

Query Match          13.2%; Score 226; DB 4; Length 382;
Best Local Similarity 26.6%; Pred. No. 1.7e-16;
Matches 63; Conservative 44; Mismatches 94; Indels 36; Gaps 5;

Qy 1 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
Db 14 PCRVLLNPOGGKGKALQLQFSRVQPFLEAEITFKLI-----LTERKNHARE 62

Qy 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120
Db 63 LVCAEELGHWDAVMGSDGLMHVVGMLMERPDWETAIOK-----PLCSLPG 110

Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLIG 171
Db 111 GSGNALAASVNHVAGYEQVTNEDLLINCTLLCRRRLSPMNLLSLHTASGLRLYSVLSLS 170

Qy 172 YGFYGDIIKDKSEKRWLGARYDFSGLKTFLSHHCYEGTVSFFLPAQHTVGSPPDRKP 228
Db 171 WGFVADVLESEKRYRLGEIRFTVGTFFRLASLRIOQLAYLP-----VGTVASKRP 223

Search completed: September 3, 2005, 04:10:35
Job time : 19.8442 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:09:18 ; Search time 66.5726 Seconds
(without alignments)
1893.930 Million cell updates/sec

Title: US-10-631-958-2
Perfect score: 1717
Sequence: 1 PHLIVFNPFGKCGKRI.....KSRNFRLRIHTNQDQ 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1717	100.0	326	5	ABb07854 Human sph
2	1717	100.0	471	4	Aae07884 Human sph
3	1640.5	95.5	416	6	ABr56301 Human sph
4	1640.5	95.5	460	3	Aay96059 Human sph
5	1640.5	95.5	481	6	ABr56302 Human sph
6	1640.5	95.5	537	5	ABb07856 Human sph
7	1640.5	95.5	537	6	Ada05680 Human NOV
8	1640.5	95.5	537	8	Adj96664 Human lip
9	1640.5	95.5	537	8	Adm62845 Human NOV
10	1640.5	95.5	562	5	ABb07857 Human sph
11	1632.5	95.1	537	5	Aam49115 Human cer
12	1503.5	87.6	531	8	Adp55248 Human PRO
13	1414.5	82.4	536	8	Adsl1054 Human the
14	1400.5	81.6	746	4	ABgl3541 Novel hum
15	1323.5	77.1	727	4	ABgl3543 Novel hum
16	1055	61.4	228	3	Aab41822 Human ORF
17	966	56.3	454	4	ABgl3544 Novel hum
18	664	38.7	136	4	ABgl3540 Novel hum
19	628	36.6	136	4	ABgl3542 Novel hum
20	613	35.7	596	4	ABb69669 Drosophil
21	476.5	27.8	182	3	Aab42383 Human ORF
22	442	25.7	532	8	Ado88891 Novel hum
23	303.5	17.7	255	8	Adsl1267 Human the
24	300.5	17.5	190	5	ABp64913 Human pro
25	282	16.4	490	6	ABr82389 D. melano

26	282	16.4	641	4	ABb58465	Drosophil
27	282	16.4	641	6	ABr82392	D. melano
28	273	15.9	144	4	AAe07885	Partial r
29	270.5	15.8	524	6	ABr82390	D. melano
30	270.5	15.8	907	4	ABb57980	Drosophil
31	270.5	15.8	907	6	ABr82393	D. melano
32	249	14.5	618	4	AAu09075	Human sph
33	243	14.2	618	4	AAU09074	Human sph
34	243	14.2	618	5	ABg31587	Human sph
35	243	14.2	618	5	ABb07855	Protein s
36	243	14.2	618	7	ADe38417	Human pro
37	243	14.2	618	8	ADQ15182	Human can
38	243	14.2	654	4	ABU52806	Human sig
39	243	14.2	654	8	ADJ66569	Sphingosi
40	243	14.2	806	4	ABG21144	Novel hum
41	236	13.7	617	4	AAU09073	Human sph
42	233.5	13.6	384	3	AAy96057	Human sph
43	228	13.3	384	5	ABb08089	Human sph
44	228	13.3	384	8	ADQ17905	Human sof
45	227.5	13.2	384	3	AAb18659	A human r

ALIGNMENTS

RESULT 1

ABb07854
ID ABb07854 standard; protein; 326 AA.

XX AC ABb07854;

XX AC ABb07854;

XX 03-JUL-2002 (first entry)

XX DE Human sphingosine kinase-like protein.

XX KW Human sphingosine kinase-like protein; intracellular signalling;

XX KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;

XX KW autoimmune disease; rheumatoid arthritis; Parkinson's disease.

XX OS Homo sapiens.

XX PN WO200228906-A2.

XX PD 11-APR-2002.

XX PF 05-OCT-2001; 2001WO-EP011516.

XX PR 06-OCT-2000; 2000US-0238005P.

XX PR 23-AUG-2001; 2001US-0314113P.

XX (FARB) BAYER AG.

XX Kossida S, Encinas J;

XX WPI; 2002-340094/37.

XX DR N-PSDB; ABL40822.

XX New reagent for modulating the activity of sphingosine kinase-like protein polypeptide or polynucleotide and treating cancer, asthma, allergy, an autoimmune disease, or a central or peripheral nervous system disorder.
XX Claim 25; Fig 2; 120pp; English.
XX The invention relates to a human sphingosine kinase-like protein. The polypeptide can be expressed by standard recombinant methodology. The sphingosine kinase-like protein and gene can be used to regulate intracellular signalling and consequently cell proliferation and apoptosis. Such regulation is useful for treating cancer, allergies (e.g. asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and peripheral nervous system disorders (e.g. Parkinson's disease). The present sequence represents the human sphingosine kinase-like protein

CC	ovarian, breast, lung, colon, testicular, stomach and skin.
CC	atherosclerosis, restenosis or ischaemia and cell proliferative disease
CC	or disorder associated with vascular diseases. Sphk gene is used in gene
CC	therapy and antisense-therapy. Sphingolipids serving as signalling
CC	molecules, have recently emerged as regulators of cell growth,
CC	differentiation, diverse cell phenotypes and cell death. Activation of
CC	SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
CC	endothelial cells. The present sequence is human sphingosine kinase
CC	(SphK) protein
XX	
SQ	Sequence 471 AA;
	Query Match 100.0%; Score 1717; DB 4; Length 471;
	Best Local Similarity 100.0%; Pred. No. 2.3e-172; Indels 0; Gaps 0
	Matches 326; Conservative 0; Mismatches 0;
Qy	1 PKHLLVFINPFGKGQKRIYERKVAPLFTLASITTDIIIGNKFYVNVYVEVITEHANQAKE 60
Db	1 PKHLLVFINPFGKGQKRIYERKVAPLFTLASITTDIIIGNKFYVNVYVEVITEHANQAKE 60
Qy	61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQSAGVDQNHPRAVLPVSSLRIGIIPA 120
Db	61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQSAGVDQNHPRAVLPVSSLRIGIIPA 120
Qy	121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 180
Db	121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 180
Qy	181 DSEKKRWGLIARYDFSGLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCROSK 240
Db	181 DSEKKRWGLIARYDFSGLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCROSK 240
Qy	241 QOLEEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
Db	241 QOLEEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
Qy	301 DLILIRKCSRFLRLIRHTNQDDQ 326
Db	301 DLILIRKCSRFLRLIRHTNQDDQ 326
RESULT 2	
AAE07884	
ID	AAE07884 standard; protein; 471 AA.
XX	
AC	AAE07884;
XX	
DT	01-NOV-2001 (first entry)
XX	
DE	Human sphingosine kinase (SphK) protein #2.
XX	
KW	Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy;
KW	antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
KW	cytostatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
KW	leukaemia; vasotropic; cell proliferative disorder; vascular disease.
XX	
OS	Homo sapiens.
XX	
PN	WO200160990-A2.
XX	
PD	23-AUG-2001.
XX	
PF	14-FEB-2001; 2001WO-US004789.
XX	
PR	14-FEB-2000; 2000US-0182360P.
XX	
PR	22-MAR-2000; 2000US-0191261P.
XX	
XX	(CURA-) CURAGEN CORP.
PA	(GETH) GENENTECH INC.
XX	
PI	Rastelli L;
XX	
DR	WPI; 2001-514770/56.
XX	
DR	N-PSDB; AAD14426.
XX	
XX	
PT	An isolated Sphingosine kinase polypeptide useful for treating a SphK-
XX	associated disorder especially cancer, restenosis or ischemia in a human.
XX	
PS	Claim 1; Fig 1; 107pp; English.
XX	
XX	The present invention relates to sphingosine kinase (SphK) polypeptides
CC	and nucleic acids encoding them. SphK is useful for treating a SphK-
CC	associated disorder especially cancers such as leukaemia, lymphoma.
XX	
PT	Platelet derived polypeptides with sphingosine kinase activity for

PT treatment of sphingosine related disorders.
 XX Claim 2; Page 29-30; 39pp; Japanese.
 XX The present sequence is the protein sequence for human sphingosine kinase
 CC 4 (SPHK4). The kinase can be used for the diagnosis and treatment of
 CC sphingosine related disorders. The kinase can also be potentially used
 CC for controlling toxicity of platelet transfusion and as a platelet
 CC stabiliser
 XX Sequence 416 AA;
 SQ Query Match 95.5%; Score 1640.5; DB 6; Length 416;
 Best Local Similarity 96.3%; Pred. No. 2.4e-164; Mismatches 0; Indels 11; Gaps 1;
 Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
 QY 1 PKHLVFINPFGGQGGKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANQAKE 60
 DB 9 PKHLVFINPFGGQGGKRIYERKVAFLFTLASITTDII-----VTEHANQAKE 57
 QY 61 TLYEINIDKYDGI VCVGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPA 120
 DB 58 TLYEINIDKYDGI VCVGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPA 117
 QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVHNSLRLRYSVSLGVGYFDIILK 180
 DB 118 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVHNSLRLRYSVSLGVGYFDIILK 177
 QY 181 DSEKKRWLGLARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCROSK 240
 DB 178 DSEKKRWLGLARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCROSK 237
 QY 241 QOLEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
 DB 238 QOLEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 297
 QY 301 DLILIRKCSRFNLFRLIRHTNQDDQ 326
 DB 298 DLILIRKCSRFNLFRLIRHTNQDDQ 323
 RESULT 4
 AAY96059
 ID AAY96059 standard; protein; 460 AA.
 AC AAY96059;
 XX 05-DEC-2000 (first entry)
 DT Human sphingosine kinase C.
 XX Sphingosine kinase C; SKC; human; drug screening; infection;
 KW antinflammatory; antiallergic; anticancer; inflammation; allergy;
 KW cancer; therapy; diagnosis.
 XX Homo sapiens.
 OS WO200052173-A2.
 FN 08-SEP-2000.
 PD 02-MAR-2000; 2000WO-CA000223.
 XX 02-MAR-1999; 99US-0122516P.
 PR (ALLX) NPS ALLELIX CORP.
 PA Munroe D, Gupta A, Falzone GR;
 XX WPI; 2000-572185/53.
 XX N-PSDB; AAA50510.
 DR New human sphingosine kinase A, B and C polynucleotides and polypeptides
 XX PT

PT useful in e.g. chromosome and gene mapping, and detecting inflammation or
 PT disease associated with abnormal levels of sphingosine kinase expression.
 XX Disclosure; Fig 9; 81pp; English.
 XX The present sequence is that of human sphingosine kinase C (SKC), an
 CC enzyme that phosphorylates sphingosine to form sphingosine 1-phosphate.
 CC The sequence was deduced from that of a polynucleotide (see AAA50510)
 CC isolated from an HeLa cDNA library. The invention provides
 CC polynucleotides (see AAA50508-10) and polypeptides (see AAY96057-59) for
 CC the human sphingosine kinase (SK) homologues SKA, SKB and SKC. The
 CC polypeptides can be obtained using recombinant DNA methods, and host
 CC cells containing expression vectors including SK polynucleotides are used
 CC in a claimed method of screening for compounds that inhibit or activate
 CC human SK activity. Human SK specific antibodies, inhibitors, ligands or
 CC their analogues can be used as bioactive agents to treat inflammation or
 CC disease including viral, bacterial or fungal infections, allergic
 CC responses, mechanical injury associated with trauma, hereditary diseases,
 CC lymphoma or carcinoma, and other conditions with activate the genes of
 CC kidney, lung, heart, lymphoid or tissues of the nervous system
 XX Sequence 460 AA;
 SQ Query Match 95.5%; Score 1640.5; DB 3; Length 460;
 Best Local Similarity 96.3%; Pred. No. 2.8e-164; Mismatches 0; Indels 11; Gaps 1;
 Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
 QY 1 PKHLVFINPFGGQGGKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANQAKE 60
 DB 53 PKHLVFINPFGGQGGKRIYERKVAFLFTLASITTDII-----VTEHANQAKE 101
 QY 61 TLYEINIDKYDGI VCVGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPA 120
 DB 102 TLYEINIDKYDGI VCVGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPA 161
 QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVHNSLRLRYSVSLGVGYFDIILK 180
 DB 162 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVHNSLRLRYSVSLGVGYFDIILK 221
 QY 181 DSEKKRWLGLARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCROSK 240
 DB 222 DSEKKRWLGLARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCROSK 281
 QY 241 QOLEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
 DB 282 QOLEEOKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 341
 QY 301 DLILIRKCSRFNLFRLIRHTNQDDQ 326
 DB 342 DLILIRKCSRFNLFRLIRHTNQDDQ 367
 RESULT 5
 ABR56302
 ID ABR56302 standard; protein; 481 AA.
 XX ABR56302;
 AC ABR56302;
 XX 20-NOV-2003 (first entry)
 DT Human Sphingosine kinase 4-related protein.
 DE Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;
 KW platelet transfusion; platelet stabiliser.
 KW Homo sapiens.
 OS WO2003031627-A1.
 FN 17-APR-2003.
 XX 28-SEP-2001; 2001WO-JP008537.
 XX

```
PR 28-SEP-2001; 2001WO-JP008537.
XX (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.
PA (CHBI-) CHEM BIOLOGY INST.
XX
PI Igarashi Y, Kihara A;
XX
DR WPI; 2003-354917/33.
XX
XX Platelet derived polypeptides with sphingosine kinase activity for
PT treatment of sphingosine related disorders.
XX
PS Example 4; Fig 4; 39pp; Japanese.
XX
XX The present invention relates to human sphingosine kinase 4 (SPHK4;
CC AB956301). The kinase can be used for the diagnosis and treatment of
CC sphingosine related disorders. The kinase can also be potentially used
CC for controlling toxicity of platelet transusion and as a platelet
CC stabiliser. The present sequence was used to illustrate the invention
XX
SQ Sequence 481 AA;

Query Match 95.5%; Score 1640.5; DB 6; Length 481;
Best Local Similarity 96.3%; Pred. No. 3e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVYVEVITEHANOAKE 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
74 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 122
|||

Qy 61 TLYEINIDKYDGVICVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
123 TLYEINIDKYDGVICVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 182
|||

Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSVVHNSLRLRYSVSLGFGYGDIIK 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
183 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSVVHNSLRLRYSVSLGFGYGDIIK 242
|||

Qy 181 DSEKKRWGLGLARYDFSGLKTFLSHHCYEGTVSFPAQHTVGSPRDRKPCRCAGCFVCQSK 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
243 DSEKKRWGLGLARYDFSGLKTFLSHHCYEGTVSFPAQHTVGSPRDRKPCRCAGCFVCQSK 302
|||

Qy 241 QOLEEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
303 QOLEEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 362
|||

Qy 301 DLILIRKCSRFNFLRLIRHTNQDQ 326
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
363 DLILIRKCSRFNFLRLIRHTNQDQ 388
|||

RESULT 6
ABB07856
ID ABB07856 standard; protein; 537 AA.
XX
XX ABB07856;
XX
XX
XX
DT 03-JUL-2002 (first entry)
XX
XX Human sphingosine kinase-like protein.
XX
XX Human sphingosine kinase-like protein; intracellular signalling;
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease.
XX
XX Homo sapiens.
XX
XX WO200228906-A2.
XX
XX 11-APR-2002.
XX
XX 05-OCT-2001; 2001WO-EP011516.
XX

PR 06-OCT-2000; 2000US-023800SP.
PR 23-AUG-2001; 2001US-0314113P.
PA (FARB ) BAYER AG.
XX
XX Kossida S, Encinas J;
XX
XX WPI; 2002-340094/37.
XX
XX N-PSDB; ABL40828.
XX
XX New reagent for modulating the activity of sphingosine kinase-like
PT protein polypeptide or polynucleotide and treating cancer, asthma,
PT allergy, an autoimmune disease, or a central or peripheral nervous system
PT disorder.
XX
XX Claim 25; Fig 10; 120pp; English.
XX
XX The invention relates to a human sphingosine kinase-like protein. The
CC polypeptide can be expressed by standard recombinant methodology. The
CC sphingosine kinase-like protein and gene can be used to regulate
CC intracellular signalling and consequently cell proliferation and
CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
CC peripheral nervous system disorders (e.g. Parkinson's disease). The
CC present sequence represents the human sphingosine kinase-like protein
XX
SQ Sequence 537 AA;

Query Match 95.5%; Score 1640.5; DB 5; Length 537;
Best Local Similarity 96.3%; Pred. No. 3.5e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVYVEVITEHANOAKE 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
130 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178
|||

Qy 61 TLYEINIDKYDGVICVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
179 TLYEINIDKYDGVICVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 238
|||

Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSVVHNSLRLRYSVSLGFGYGDIIK 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
239 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSVVHNSLRLRYSVSLGFGYGDIIK 298
|||

Qy 181 DSEKKRWGLGLARYDFSGLKTFLSHHCYEGTVSFPAQHTVGSPRDRKPCRCAGCFVCQSK 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
299 DSEKKRWGLGLARYDFSGLKTFLSHHCYEGTVSFPAQHTVGSPRDRKPCRCAGCFVCQSK 358
|||

Qy 241 QOLEEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
359 QOLEEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418
|||

Qy 301 DLILIRKCSRFNFLRLIRHTNQDQ 326
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
419 DLILIRKCSRFNFLRLIRHTNQDQ 444
|||

RESULT 7
ADA05680
ID ADA05680 standard; protein; 537 AA.
XX
XX ADA05680;
XX
XX
XX
DT 06-NOV-2003 (first entry)
XX
XX Human NOV9a protein SEQ ID NO:40.
XX
XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
KW immunomodulator; cytostatic; nootropic; neuroprotective;
KW antiparkinsonian; antilipemic; gene therapy; human disease;
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia.
```

XX Homo sapiens.
XX WO2003029424-A2.
XX 10-APR-2003.
XX 02-OCT-2002; 2002WO-US031373.
XX 02-OCT-2001; 2001US-0326483P.
XX 05-OCT-2001; 2001US-0327435P.
XX 05-OCT-2001; 2001US-0327449P.
XX 09-OCT-2001; 2001US-0327917P.
XX 09-OCT-2001; 2001US-0328029P.
XX 09-OCT-2001; 2001US-0328044P.
XX 09-OCT-2001; 2001US-0328056P.
XX 12-OCT-2001; 2001US-0328849P.
XX 15-OCT-2001; 2001US-0329414P.
XX 17-OCT-2001; 2001US-0330142P.
XX 18-OCT-2001; 2001US-0330309P.
XX 22-OCT-2001; 2001US-0341058P.
XX 24-OCT-2001; 2001US-0339266P.
XX 29-OCT-2001; 2001US-0343629P.
XX 29-OCT-2001; 2001US-0349575P.
XX 01-NOV-2001; 2001US-0346357P.
XX 17-APR-2002; 2002US-0373260P.
XX 19-APR-2002; 2002US-0373815P.
XX 19-APR-2002; 2002US-0373817P.
XX 19-APR-2002; 2002US-0373826P.
XX 19-APR-2002; 2002US-0373884P.
XX 22-APR-2002; 2002US-0374977P.
XX 16-MAY-2002; 2002US-0381037P.
XX 16-MAY-2002; 2002US-0381038P.
XX 16-MAY-2002; 2002US-0381042P.
XX 17-MAY-2002; 2002US-0381642P.
XX 28-MAY-2002; 2002US-0383656P.
XX 29-MAY-2002; 2002US-0383831P.
XX 25-JUN-2002; 2002US-0391335P.
XX 01-OCT-2002; 2002US-00262511.
XX (CURA-) CURAGEN CORP.
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patcurajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI Ort T, Gorman L, Zehrhusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
XX WPI; 2003-381626/36.
DR N-PSDB; ADA05679.
XX
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Claim 1; Page 135; 586pp; English.
XX
XX The present invention describes NOVX proteins, where X can be 1 to 55
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
CC described above and a carrier; (2) a kit comprising, in one or more
CC containers, the composition described above; (3) an isolated nucleic acid
CC molecule which encodes a NOVX protein of the invention; (4) a vector
CC comprising the nucleic acid molecule described above; (5) a cell
CC comprising the above vector; (6) an antibody that immunospecifically
CC binds to the polypeptide described above; (7) methods for determining the
CC presence or amount of the above polypeptide or nucleic acid molecule in a
CC sample; (8) methods for determining the presence of or predisposition to
CC a disease associated with altered levels of expression of the above
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC method of identifying an agent that binds to the polypeptide described
CC above; (10) a method for identifying a potential therapeutic agent for

CC use in treating a pathology that is related to an aberrant expression or
CC aberrant physiological interactions of the polypeptide; (11) a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide; (12) a method for modulating
CC the activity of the polypeptide described above; (13) methods of treating
CC or preventing a pathology associated with the above polypeptide in a
CC mammal; and (14) a method for producing the above polypeptide. NOVX
CC sequences have antidiabetic, anorectic, antibacterial, virucide,
CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
CC and antilipemic activities, and can be used in gene therapy. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease. The polypeptide or the nucleic
CC acid molecule may be used to diagnose, treat or prevent metabolic
CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence represents a human NOVX from the
CC present invention.
XX
XX Sequence 537 AA;
SQ
Query Match 95.5%; Score 1640.5; DB 6; Length 537;
Best Local Similarity 96.3%; Pred No. 3.5e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
QY 1 PKHLVFINPFGGKQGRKRIYERKVPFLTLASITTDIIIGNKFYVYVEVITEHANQAKE 60
DB 130 PKHLVFINPFGGKQGRKRIYERKVPFLTLASITTDII-----VTEHANQAKE 178
QY 61 TLYEINIDKYGIVCVGDDGMFSEVLHGLIGRTQBSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 179 TLYEINIDKYGIVCVGDDGMFSEVLHGLIGRTQBSAGVDQNHPRAVLVPSSLRIGIIPA 238
QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHNSTLLRYSVSLGYGFGYDIIK 180
DB 239 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHNSTLLRYSVSLGYGFGYDIIK 298
QY 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTGSFPAQHTVGSPPDRKPCACGFCVCRQSK 240
DB 299 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTGSFPAQHTVGSPPDRKPCACGFCVCRQSK 358
QY 241 QOLEBEQKALYGLEAAEDVEEWQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 300
DB 359 QOLEBEQKALYGLEAAEDVEEWQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 418
QY 301 DLILIRKCSRFNLFRLIRHTNQDDQ 326
DB 419 DLILIRKCSRFNLFRLIRHTNQDDQ 444

RESULT 8
ADJ96664
ID ADJ96664 standard; protein; 537 AA.
XX
XX AC ADJ96664;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE Human lipid kinase KIAA1646 protein SeqID 121.
XX
XX KW kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
XX PTK; STK; gene therapy; cancer; immune-related disease;
XX cardiovascular disease; brain; neuronal associated disease; metabolic;
XX inflammatory disorder; cytostatic; neuroprotective; immunomodulator;
XX antinflammatory; enzyme; lipid kinase; KIAA1646.
XX
XX OS Homo sapiens.
XX 44.
XX
XX PN WO2004006838-A2.
XX

PD 22-JAN-2004.
XX
PF 15-JUL-2003; 2003WO-US021730.
XX
PR 15-JUL-2002; 2002US-0395632P.
XX
XX (SUGE-) SUGEN INC.
XX
XX Whyte D, Manning G, Caenepeel S;
XX
XX WPI; 2004-122753/12.
DR N-PSDB; ADJ96598.
XX
XX New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
PS Claim 1; SEQ ID NO 121; 366pp; English.
XX
XX This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytosolic, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polypeptide sequence is a human kinase protein sequence
CC of the invention.
XX
XX Sequence 537 AA;
SQ
Query Match 95.5%; Score 1640.5; DB 8; Length 537;
Best Local Similarity 96.3%; Pred. No. 3.5e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
QY 1 PKHLVFNPPGGKGGRRIYERKVPFLFTLASITTDIIGNKFYVNVVEVITEHANOAKE 60
DB 130 PKHLVFNPPGGKGGRRIYERKVPFLFTLASITTDII-----VTEHANOAKE 178
QY 61 TLYEINIDYDGI VCVGSDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 179 TLYEINIDYDGI VCVGSDGMFSEVLHGLIGRTQESAGVDQNHPRAVLVPSSLRIGIIPA 238
QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSVHNSHTLLRYSVSLGCGFYGDIK 180
DB 239 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSVHNSHTLLRYSVSLGCGFYGDIK 298
QY 181 DSEKKRWGLARYDFSGIKTFLSHHCYGTGVSFLPAQHTVCSPRDRKPCRAGCFVCRQSK 240
DB 299 DSEKKRWGLARYDFSGIKTFLSHHCYGTGVSFLPAQHTVCSPRDRKPCRAGCFVCRQSK 358
QY 241 QOLBEEQKALYGLEAEDVEWQVCGKFLAINATNNSCARRSPRGLSPAHLGDGSS 300
DB 359 QOLBEEQKALYGLEAEDVEWQVCGKFLAINATNNSCARRSPRGLSPAHLGDGSS 418
QY 301 DLILIRKCSRFNFLRLIRHTNQDQ 326
DB 419 DLILIRKCSRFNFLRLIRHTNQDQ 444
RESULT 9
ADN62845
ID ADN62845 standard; protein; 537 AA.
XX
XX AC ADN62845;
XX
XX DT 01-JUL-2004 (first entry)
XX
XX DE Human NOV9a.

XX human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;
KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;
wasting disorder.
XX
XX Homo sapiens.
OS
XX US2004038223-A1.
PN
XX 26-FEB-2004.
PD
XX
XX 01-OCT-2002; 2002US-00262511.
XX
XX 02-OCT-2001; 2001US-0326483P.
PR
XX 05-OCT-2001; 2001US-0327435P.
PR
XX 05-OCT-2001; 2001US-0327449P.
PR
XX 09-OCT-2001; 2001US-0327917P.
PR
XX 09-OCT-2001; 2001US-0328029P.
PR
XX 09-OCT-2001; 2001US-0328044P.
PR
XX 09-OCT-2001; 2001US-0328056P.
PR
XX 12-OCT-2001; 2001US-0328849P.
PR
XX 15-OCT-2001; 2001US-0329414P.
PR
XX 17-OCT-2001; 2001US-0330142P.
PR
XX 18-OCT-2001; 2001US-0330309P.
PR
XX 22-OCT-2001; 2001US-0341058P.
PR
XX 24-OCT-2001; 2001US-0339266P.
PR
XX 24-OCT-2001; 2001US-0343629P.
PR
XX 29-OCT-2001; 2001US-0349575P.
PR
XX 01-NOV-2001; 2001US-0346357P.
PR
XX 17-APR-2002; 2002US-0373260P.
PR
XX 19-APR-2002; 2002US-0373815P.
PR
XX 19-APR-2002; 2002US-0373817P.
PR
XX 19-APR-2002; 2002US-0373826P.
PR
XX 19-APR-2002; 2002US-0373884P.
PR
XX 22-APR-2002; 2002US-0374977P.
PR
XX 16-MAY-2002; 2002US-0381037P.
PR
XX 16-MAY-2002; 2002US-0381038P.
PR
XX 16-MAY-2002; 2002US-0381042P.
PR
XX 17-MAY-2002; 2002US-0381642P.
PR
XX 28-MAY-2002; 2002US-0383656P.
PR
XX 29-MAY-2002; 2002US-0383831P.
PR
XX 25-JUN-2002; 2002US-0391335P.
XX
XX (SMIT/) SMITHSON G.
PA (MILL/) MILLET I.
PA (PEYM/) PEYMAN J A.
PA (KEKU/) KEKUDA R.
PA (JUJJ/) JU J.
PA (LILL/) LI L.
PA (GUOX/) GUO X.
PA (PATT/) PATTURAJAN M.
PA (SPYT/) SPYTEK K A.
PA (EDIN/) EDINGER S R.
PA (ELLE/) ELLERMAN K.
PA (MALY/) MALYANKAR U M.
PA (ORTT/) ORT T.
PA (GORM/) GORMAN L.
PA (ZERH/) ZERHUSEN B D.
PA (ANDE/) ANDERSON D W.
PA (ZHON/) ZHONG M.
PA (CATT/) CATTERTON E.
PA (JIWW/) JI W.
PA (MILL/) MILLER C E.
PA (RAST/) RASTELLI L.
PA (STON/) STONE D J.
PA (PENA/) PENNA C E A.
PA (SHEN/) SHENOY S G.
PA (SHIM/) SHIMKETS R A.
PA (ROTH/) ROTHENBERG M E.
PA (LEAC/) LEACH M D.
PA (AGEE/) AGEE M L.

```
PA (BERG/) BERGHS C.
PA (DIP/) DIPPO V A.
PA (EISE/) EISEN A.
PA (GANG/) GANGOLLI E A.
PA (RIEG/) RIEGER D K.
PA (SPAD/) SPADERNA S K.
XX
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;
XX
DR WPI; 2004-213931/20.
DR N-PSDB; ADN62844.
XX
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.
XX
XX Claim 1; SEQ ID NO 40; 395pp; English.
XX
PS The invention relates to isolated NOVX polypeptides and polynucleotides.
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or
CC treat a medical condition in human related to the aberrant expression and
CC activity of NOVX polypeptides. For example, NOVX polypeptides and
CC polynucleotides may be used to treat disorders associated with decreased
CC expression or activity of NOVX by supplementing the patient our
CC production or to rectify mutations. Conversely, antisense NA molecules
CC may be administered to down regulate expression of NOVX polypeptides by
CC binding with the cells own genes and preventing their expression. NOVX
CC polynucleotides and complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar sequences in samples, and so which patients may be in need of
CC restorative therapy. NOVX polypeptides may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of NOVX. The
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be
CC used to modulate NOVX polynucleotide expression and activity of NOVX
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX
CC polypeptides and polynucleotides may be used in this way to prevent,
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
CC haematopoietic disorders, and the various dyslipidaemias, metabolic
CC disturbances associated with obesity, the metabolic syndrome X and
CC wasting disorders associated with chronic diseases and various cancers.
CC They may also be used as antibacterial agents. The present sequence
CC represents the amino acid sequence of a human NOVX protein.
XX
SQ Sequence 537 AA;
Query Match 95.5%; Score 1640.5; DB 8; Length 537;
Best Local Similarity 96.3%; Pred. No. 3.5e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
QY 1 PKHLVFNPFPGKGQGRKRIYERKVAFLFTLASITTDIIIGNKFYVYVVEVITEHANOAKE 60
Db 130 PKHLVFNPFPGKGQGRKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178
QY 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPA 120
Db 179 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPA 238
QY 121 GSTDCVCYVTGTSDAETSAHIVVGDLSLMDVSSVHHNSTLLRYSVLLGFGFYGDIIK 180
Db 239 GSTDCVCYVTGTSDAETSAHIVVGDLSLMDVSSVHHNSTLLRYSVLLGFGFYGDIIK 298
QY 181 DSEKKRWLGLARYDFSLGKTLFSLHHCYEGTGVSLFPAQHTVGSPRDKPCRCACFCVCROSK 240
Db 299 DSEKKRWLGLARYDFSLGKTLFSLHHCYEGTGVSLFPAQHTVGSPRDKPCRCACFCVCROSK 358
```

```
QY 241 QOLBEEQKALYGLLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLIGDSS 300
Db 359 QOLBEEQKALYGLLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLIGDSS 418
QY 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326
Db 419 DLILIRKCSRFNFLRLIRHTNQDDQ 444
RESULT 10
ABB07857
ID ABB07857 standard; protein; 562 AA.
XX
AC ABB07857;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human sphingosine kinase-like protein.
XX
KW Human sphingosine kinase-like protein; intracellular signalling;
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease.
XX
OS Homo sapiens.
XX
PN WO200228906-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-EP011516.
XX
PR 06-OCT-2000; 2000US-0238005P.
PR 23-AUG-2001; 2001US-0314113P.
XX
PA (FARB ) BAYER AG.
XX
PI Kossida S, Encinas J;
XX
WPI; 2002-340094/37.
N-PSDB; ABL40828.
XX
XX New reagent for modulating the activity of sphingosine kinase-like
XX protein polypeptide or polynucleotide and treating cancer, asthma,
XX allergy, an autoimmune disease, or a central or peripheral nervous system
XX disorder.
PS Claim 25; Fig 11; 120pp; English.
XX
CC The invention relates to a human sphingosine kinase-like protein. The
CC polypeptide can be expressed by standard recombinant methodology. The
CC sphingosine kinase-like protein and gene can be used to regulate
CC intracellular signalling and consequently cell proliferation and
CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
CC peripheral nervous system disorders (e.g. Parkinson's disease). The
CC present sequence represents the human sphingosine kinase-like protein
XX
SQ Sequence 562 AA;
Query Match 95.5%; Score 1640.5; DB 5; Length 562;
Best Local Similarity 96.3%; Pred. No. 3.7e-164;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
QY 1 PKHLVFNPFPGKGQGRKRIYERKVAFLFTLASITTDIIIGNKFYVYVVEVITEHANOAKE 60
Db 155 PKHLVFNPFPGKGQGRKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 203
QY 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPA 120
Db 204 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPA 263
QY 121 GSTDCVCYVTGTSDAETSAHIVVGDLSLMDVSSVHHNSTLLRYSVLLGFGFYGDIIK 180
```

Db 264 GSTDCVSVTGTSDAETSAHIVVGDGLAMDVSSVHHNSTLLRYSVSLGFGYFDIHK 323
Qy 181 DSEKKRWGLARYDFSGDKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGCFVCRQSK 240
Db 324 DSEKKRWGLARYDFSGDKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGCFVCRQSK 383
Qy 241 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMNSCACRRSPRGLSPAHLGDGSS 300
Db 384 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMNSCACRRSPRGLSPAHLGDGSS 443
Qy 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326
Db 444 DLILIRKCSRFNFLRLIRHTNQDDQ 469
RESULT 11
AAW49115
ID AAW49115 standard; protein; 537 AA.
XX AC AAW49115;
XX DT 20-MAY-2002 (first entry)
XX DE Human ceramide kinase hCERK1.
XX KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;
KW neurological disease; inflammation; human immunodeficiency virus;
KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;
KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;
KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;
KW enzyme.
XX OS Homo sapiens.
XX PN WO200196575-A1.
XX PD 20-DEC-2001.
XX PF 11-JUN-2001; 2001WO-JP004889.
XX PR 14-JUN-2000; 2000JP-00178039.
XX PA (SANY) SANKYO CO LTD.
XX PI Sugiyura M, Kono K, Kohama T;
XX WPI; 2002-179513/23.
XX DR N-PSDB; ABA96945.
XX PT Human ceramide kinase gene and the enzyme encoded by it for screening
PT substances as drugs for neurological, inflammatory and other disorders.
XX PS Claim 1; Page 54-57; 61pp; Japanese.
XX CC This sequence represents a human ceramide kinase designated hCERK1. The
CC invention relates to hCERK1, nucleic acids encoding it, expression
CC vectors and host cells containing hCERK1 nucleic acids, the recombinant
CC production of hCERK1 and antibodies specific for hCERK1. The invention
CC also encompasses methods of isolating hCERK1 from samples, the use of
CC hCERK1 in drug screening, and the use of hCERK1 nucleic acid sequences in
CC gene therapy. hCERK1 mediates the ATP-dependent 1-phosphorylation of
CC ceramides and can be used to screen for therapeutic and preventive agents
CC for a wide range of disorders. Such disorders include neurological
CC disease, inflammation, human immunodeficiency virus (HIV) infection, type
CC 2 diabetes, obesity, sepsis, arteriosclerosis and cancer
XX SQ Sequence 537 AA;
Query Match 95.1%; Score 1632.5; DB 5; Length 537;
Best Local Similarity 95.7%; Pred No. 2.5e-163;
Matches 312; Conservative 1; Mismatches 1; Indels 11; Gaps 1;
Qy 1 PKHLLVFINFPGKGQKRIYERKVAPLFTLASITTDII 1

Db 130 PKHLLVFINFPGKGQKRIYERKVAPLFTLASITTDII 178
Qy 61 TLYEINIDKDYGVVCVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
Db 179 TLYEINIDKDYGVVCVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 238
Qy 121 GSTDCVSVTGTSDAETSAHIVVGDGLAMDVSSVHHNSTLLRYSVSLGFGYFDIHK 180
Db 239 GPTDCVSVTGTSDAETSAHIVVGDGLAMDVSSVHHNSTLLRYSVSLGFGYFDIHK 298
Qy 181 DSEKKRWGLARYDFSGDKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGCFVCRQSK 240
Db 299 DSEKKRWGLARYDFSGDKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGCFVCRQSK 358
Qy 241 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMNSCACRRSPRGLSPAHLGDGSS 300
Db 359 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMNSCACRRSPRGLSPAHLGDGSS 418
Qy 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326
Db 419 DLILIRKCSRFNFLRLIRHTNQDDQ 444
RESULT 12
ADP55248
ID ADP55248 standard; protein; 531 AA.
XX AC ADP55248;
XX DT 18-NOV-2004 (first entry)
XX DE Human PRO protein sequence SEQ ID NO:1224.
XX KW human; PRO; immune related disease; inflammatory immune response;
KW immune response stimulation; antiallergic; antianaemic; antiarthritic;
KW antiaesthetic; antidiabetic; antiinflammatory; antipsoriatic;
KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW virucide; gene therapy.
XX OS Homo sapiens.
XX PN WO2004039956-A2.
XX PD 13-MAY-2004.
XX PF 28-OCT-2003; 2003WO-US034381.
XX PR 29-OCT-2002; 2002US-0422472P.
XX PA (GETH) GENENTECH INC.
XX PI Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX DR WPI; 2004-376182/35.
XX DR N-PSDB; ADP55247.
XX CC New PRO polynucleotides and polypeptides, useful in useful in diagnosing
XX and treating an immune related disease, e.g. systemic lupus
XX erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
XX stimulating an immune response.
XX PS Claim 1; SEQ ID NO 1224; 3009pp; English.
XX CC The present invention describes an isolated PRO nucleic acid (1). Also
CC described: (1) a vector comprising (1); (2) a host cell comprising the
CC vector of (1); (3) a process for producing a PRO polypeptide; (4) an
CC isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC antibody which specifically binds to a polypeptide of (4); (7) a


```

OS Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS77728.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 20; SEQ ID NO 43900; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 746 AA;
XX Query Match 81.6%; Score 1400.5; DB 4; Length 746;
XX Best Local Similarity 80.8%; Pred. No. 1.4e-138;
XX Matches 274; Conservative 4; Mismatches 0; Indels 61; Gaps 2;

Qy 49 EVITEHANQAKETLYEINIDKYGVGCGMFSEVLHGLIGRTQRSAGVDQNHPRAVL 108
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
298 KLITEHANQAKETLYEINIDKYGVGCGMFSEVLHGLIGRTQRSAGVDQNHPRAVL 357
Qy 109 VPSSLRIGIIPA-----GSTDVCYSTVGTSDAETSAIHIV----- 145
Db 358 VPSSLRIGIIPAHHVGPGENAGGLDGDTCVCYSTVGTSDAETSAIHIVGCCPEARKPP 417
Qy 146 -----GDSLAMDVSSVHNSLTLRYSV 167
Db 418 ASRHATCGSHGQLCLGCOLNACLCEASRLQSRMQSGDSLAMDVSSVHNSLTLRYSV 477
Qy 168 SLLGYGFYDIIKQSEKKRWLGARYDFSLKTLFSLHHCYEGTGVSFPAQHTVGSPRDRK 227
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
478 SLLGYGFYDIIKQSEKKRWLGARYDFSGIKTLFSLHHCYEGTGVSFPAQHTVGSPRNRK 537
Qy 228 PCRACGFCVCRSKQQLBEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPR 287
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
538 PCRACGFCVCRSKQQLBEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPR 597

```

```

Qy 288 GLSPAHLGDSSDLILIRKCSRPNFLRLIRHTNQDQ 326
Db |||||
598 GLSPAHLGDSSDLILIRKCSRPNFLRLIRHTNQDQ 636

RESULT 15
ABG13543
ID ABG13543 standard; protein; 727 AA.
XX
XX ABG13543;
XX
XX 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #13534.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS77730.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 20; SEQ ID NO 43902; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 727 AA;
XX Query Match 77.1%; Score 1323.5; DB 4; Length 727;
XX Best Local Similarity 58.3%; Pred. No. 2e-130;
XX Matches 275; Conservative 0; Mismatches 0; Indels 197; Gaps 3;

Qy 52 TEHANQAKETLYEINIDKYDG-----
Db |||||

```


Db 1 TEHANQAKETLYEINIDKYDGYTAHSRMPCCPRNHSAGKAPVQGVGEGQRAFPWEP 60
Qy 73 -----IVCVGGDGMFSEVLHGL 89
Db 61 GPKOMPAKQVCRLEPLKLEASGLLRSEBRTCRSPDRLCSCSIVCVGGDGMFSEVLHGL 120
Qy 90 IGRORSAGVDONHPRAVLVPSSLRIGIIPA----- 120
Db 121 IGRORSAGVDONHPRAVLVPSSLRIGIIPAGFANDLAGDEVSVLSGPVSGPEGGMVHT 180
Qy 121 -----GSTDCVCYSTVGTSDAETSALH 142
Db 181 FTLVTALGCEHRSHPHFMDERRTGEHVPGENAGGLDSTDCVCYSTVGTSDAETSALH 240
Qy 143 IVW-----GDSLAMDVS 154
Db 241 IVWGCPPEARPPASRHTACSGHGQCLGCOLMNACLEASRLQSRMQSPGDSLAMDV 300
Qy 155 SVHHNSTLLRYSVSLGFGYDIIKDSEKKRWLGARYDFGLKTFLSHHCYEGTVSFL 214
Db 301 SVHHNSTLLRYSVSLGFGYDIIKDSEKKRWLGARYDFGLKTFLSHHCYEGTVSFL 360
Qy 215 PAQHTVGSPRDRKPCRCAGCFVCRQSKQLEBEQKALYGLEAAEDVEEWQVVCCKFLAIN 274
Db 361 PAQHTVGSPRDRKPCRCAGCFVCRQSKQLEBEQKALYGLEAAEDVEEWQVVCCKFLAIN 420
Qy 275 ATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNRLFLIRHTNQDDO 326
Db 421 ATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNRLFLIRHTNQDDO 472

Search completed: September 3, 2005, 04:37:38
Job time : 74.5726 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:25:44 ; Search time 104.512 Seconds
(without alignments)
2753.633 Million cell updates/sec

Title: US-10-631-958-11

Perfect score: 3025

Sequence: 1 HEAANGPAPLGVAPPARWT.....QLVRLFARGIENPKPDSHS 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	2888	95.5	537	1 CEK1_HUMAN	Q8tct0 homo sapien
2	2428	80.3	531	1 CEK1_MOUSE	Q8k4q7 mus musculu
3	1904	62.9	409	2 Q62PK5	Q6zpk5 mus musculu
4	1811	59.9	339	2 Q6NKS5	Q6nxs5 homo sapien
5	1713	56.6	572	2 Q6GLV1	Q6glv1 xenopus lae
6	1495.5	49.4	485	2 Q6GMF3	Q6gmf3 xenopus lae
7	801.5	26.5	687	2 Q9VNA6	Q9vna6 drosophila
8	763	25.2	487	2 Q9S5L5	Q9s5l5 drosophila
9	666.5	22.0	410	2 Q7PRA8	Q7pra8 anopheles g
10	648	21.4	608	2 Q6USK2	Q6usk2 arabidopsis
11	620.5	20.5	700	2 Q6HGK1	Q6hgh1 oryza sativ
12	594	19.6	532	2 Q6UZF6	Q6uzf6 homo sapien
13	583.5	18.3	533	2 Q9LU45	Q9lu45 arabidopsis
14	567.5	18.8	549	2 Q9T2I1	Q9t2i1 caenorhabdi
15	480	15.9	586	2 Q949C3	Q949c3 oryza sativ
16	428	14.1	485	2 Q8I7L1	Q8i7l1 arabidopsis
17	417.5	13.8	1240	2 Q65419	Q65419 arabidopsis
18	373.5	12.3	480	2 Q8H350	Q8h350 oryza sativ
19	370.5	12.2	763	2 Q9LRB0	Q9lrb0 arabidopsis
20	350	11.6	732	2 Q9PHL3	Q9phl3 arabidopsis
21	343	11.3	280	2 Q6ZP59	Q6zps9 homo sapien
22	331.5	11.0	446	2 Q84S01	Q84s01 oryza sativ
23	330.5	10.9	757	2 Q94HY9	Q94hy9 oryza sativ
24	330.5	10.9	757	2 Q7XCS9	Q7xcs9 oryza sativ
25	323	10.7	685	2 Q7QIP4	Q7qip4 anopheles g
26	314.5	10.4	748	2 Q7XN57	Q7xn57 oryza sativ
27	310.5	10.3	641	2 Q9VYI8	Q9vyi8 drosophila
28	298.5	9.9	661	2 Q9VZW0	Q9vzw0 drosophila
29	291.5	9.6	458	2 Q14159	Q14159 schizosacch
30	277	9.2	654	1 SPH2_HUMAN	Q9nia7 mus musculu
31	267.5	8.8	616	2 Q6AYB2	Q6ayb2 rattus norv

32	267.5	8.8	517	1 SPH2_MOUSE	Q9nia7 mus musculu
33	264.5	8.7	384	1 SPH1_HUMAN	Q9nyal homo sapien
34	264.5	8.7	384	2 Q96HV8	Q96hv8 homo sapien
35	264.5	8.7	398	2 Q96GK1	Q96gk1 homo sapien
36	264.5	8.7	470	2 Q8N632	Q8n632 homo sapien
37	260	8.6	388	2 Q88886	Q88886 mus musculu
38	257.5	8.5	624	2 Q86KF9	Q86kf9 dictyosteli
39	254	8.4	381	2 Q91ZN3	Q91zn3 mus musculu
40	254	8.4	382	2 Q8C1I5	Q8c1i5 mus musculu
41	254	8.4	504	2 Q88885	Q88885 mus musculu
42	253.5	8.4	579	2 Q74ZE3	Q74ze3 ashbya goss
43	248.5	8.2	687	2 Q06147	Q06147 saccharomyc
44	245.5	8.1	383	2 Q642F6	Q642f6 rattus norv
45	245.5	8.1	423	2 Q7JM91	Q7jtm91 caenorhabdi

ALIGNMENTS

RESULT 1
CEK1_HUMAN
ID CEK1_HUMAN STANDARD; PRT; 537 AA.
AC Q8TCT0; Q9BYB3; Q9UGB5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ceramide kinase (EC 2.7.1.138) (Acylsphingosine kinase) (hCERK) (Lipid
kinase 4) (Lk4).
GN Name=CERK; Synonyms=KIAA1646;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Leukemia;
RX MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;
RA Sugiura M., Kono K., Liu H., Shimizugawa T., Minekura H., Spiegel S.,
Kohama T.;
RT "Ceramide kinase, a novel lipid kinase. Molecular cloning and
functional characterization."
RL J. Biol. Chem. 277:23294-23300(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Van Veldhoven P.P.;
RT "A search for lipid kinases."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
Bagnuley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
Laird G.K., Langford C.F., Leverhwa M.A., Lloyd C., Lloyd D.M.,
Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
Odel C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
Soderlund C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,
Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
Minoshima S., Kawasaki K., Sasaki T., Aakawa S., Kudoh J.,

RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Pang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.I.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoso S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Schaefer P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.M., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saito S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelman L., Kim U.J., Shizuya H., Simon M.I., Dumanek J.P.,
RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
RN [4]
RP SEQUENCE OF 57-537 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21156230; PubMed=11258795;
RA Hirotsawa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.;
RT "Identification of novel transcribed sequences on human chromosome 22
RT by expressed sequence tag mapping.";
RL DNA Res. 8:1-9(2001).
CC -!- FUNCTION: Catalyzes specifically the phosphorylation of ceramide
CC to form ceramide 1-phosphate. Acts efficiently on natural and
CC analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),
CC to a lesser extent on C2-ceramide and C6-dihydroceramide, but not
CC on other lipids, such as various sphingosines
CC -!- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.
CC -!- COFACTOR: Calcium and magnesium.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC -!- TISSUE SPECIFICITY: High level expression in heart, brain,
CC skeletal muscle, kidney and liver; moderate in peripheral blood
CC leukocytes and thymus; very low in spleen, small intestine,
CC placenta and lung.
CC -!- MISCELLANEOUS: Optimal pH is 6.0-7.5.
CC -!- SIMILARITY: Contains 1 DAGKC domain.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to erroneous
CC gene model prediction. An additional exon may exist between amino
CC acid positions 168 and 169.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB079066; BAC01154.1; -;
DR EMBL; AJ457828; CAD29884.1; -;
DR EMBL; AL096766; CAB62977.1; ALT SEQ.
DR EMBL; AL118516; -; NOT ANNOTATED_CDS.
DR EMBL; AB051433; BAB33316.1; -;
DR Genew; HGNC:19256; CERK.
DR GO; GO:0000299; C:integral to membrane of membrane fraction; IDA.
DR GO; GO:0004685; F:calcium/calmodulin-dependent protein kinase. . .; IDA.
DR GO; GO:0001729; F:ceramide kinase activity; IDA.
DR GO; GO:0000287; F:magnesium ion binding; IDA.
DR GO; GO:0006672; P:ceramide metabolism; TAS.
DR InterPro; IPR001206; DAGKC.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011036; PH-related.
DR Pfam; PF00781; DAGK_cat.1.
DR ProDom; PD005043; DAGKC; 1.
DR SMART; SM00046; DAGKC; 1.
DR SMART; SM00233; PH; 1.
KW Calcium; Kinase; Magnesium; Transferase.

FT DOMAIN 132 278 DAGKC.
SQ SEQUENCE 537 AA; 59977 MW; 3DBFC0ED8D679F7F CRC64;
Query Match 95.5%; Score 2888; DB 1; Length 537;
Best Local Similarity 100.0%; Pred. NO. 1.6e-227;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 26 MGATGAAPLQSVLVVVKQRCVAVSLPARALLRWRSPGCGAGAGADACVSPVSEIIAV 85
Db 1 MGATGAAPLQSVLVVVKQRCVAVSLPARALLRWRSPGCGAGAGADACVSPVSEIIAV 60
Qy 86 EETDVHGHQSGKQWKQXMEKPYAFTVHCVKRRRRWKAQVTPWCPEEQQLCHLWLOTLR 145
Db 61 EETDVHGHQSGKQWKQXMEKPYAFTVHCVKRRRRWKAQVTPWCPEEQQLCHLWLOTLR 120
Qy 146 EMLEKLTSRPKHLLVFNPPFGKGQGGKRIYERKVPALFTLASITTDIIIVTEHANAQKETT 205
Db 121 EMLEKLTSRPKHLLVFNPPFGKGQGGKRIYERKVPALFTLASITTDIIIVTEHANAQKETT 180
Qy 206 YEINIDKYDGI VCGDGMFSEVLHGLIGRTQRSAGVDQNHPRVAVLPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGI VCGDGMFSEVLHGLIGRTQRSAGVDQNHPRVAVLPSSLRIGIIPAGS 240
Qy 266 TDCVCYSTGTSDAETSGALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLLGFGFYGDIKDS 325
Db 241 TDCVCYSTGTSDAETSGALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLLGFGFYGDIKDS 300
Qy 326 EKKRWGLGARYDFSGLTFLSHHCYEGTVSGFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 385
Db 301 EKKRWGLGARYDFSGLTFLSHHCYEGTVSGFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 360
Qy 386 LEEQKALYGLEAAEDVEEWQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445
Db 361 LEEQKALYGLEAAEDVEEWQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Qy 446 ILIRKCSRNFPLIRLIRHTNQDQDFTFVEVYRVKFKQFTSKIMEDESDSLKGGKKRF 505
Db 421 ILIRKCSRNFPLIRLIRHTNQDQDFTFVEVYRVKFKQFTSKIMEDESDSLKGGKKRF 480
Qy 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPALIEVRVHCOLVRLFARGIEENPKPDSSH 562
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPALIEVRVHCOLVRLFARGIEENPKPDSSH 537
RESULT 2
CEKI_MOUSE STANDARD; PRT; 531 AA.
ID CEKI_MOUSE
AC Q8K4Q7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ceramide kinase (EC 2.7.1.138) (Acylsphingosine kinase) (mCERK).
GN Name=Cerk;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;
RA Sugiyama M., Kono K., Liu H., Shimizuogawa T., Minekura H., Spiegel S.,
RA Kohama T.;
RT "Ceramide kinase, a novel lipid kinase. Molecular cloning and
RT functional characterization.";
RL J. Biol. Chem. 277:23294-23300(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart, and Thymus;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Saito K., Nishikawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Perte G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Iehli Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
CC -!- FUNCTION: Catalyzes specifically the phosphorylation of ceramide
to form ceramides 1-phosphate. Acts efficiently on natural and
analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),
to a lesser extent on C2-ceramide and C6-dihydroceramide, but not
on other lipids, such as various sphingosines (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.
CC -!- COFACTOR: Calcium and magnesium (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
similarity).
CC -!- TISSUE SPECIFICITY: High level expression in heart, brain and
testis; low expression in spleen, liver and lung; not detected in
skeletal muscle.
CC -!- DEVELOPMENTAL STAGE: Highly expressed at 7 dpc and decreases
rapidly thereafter.
CC -!- SIMILARITY: Contains 1 DAGKc domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB079067; BAC01155.1; -;
DR EMBL; AK042077; BAC31157.1; -;
DR EMBL; AK052269; BAC34908.1; -;
DR MGD; MGI:2386052; Cerk.
DR GO; GO:0000299; C:integral to membrane of membrane fraction; ISS.
DR GO; GO:0004685; F:calcium/calmodulin-dependent protein kinase. . .; ISS.
DR GO; GO:0001729; F:ceramide kinase activity; ISS.
DR GO; GO:0000287; F:magnesium ion binding; ISS.
DR GO; GO:0006672; P:ceramide metabolism; ISS.
DR InterPro; IPR001206; DAGKc.
DR InterPro; IPR011036; PH-related.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKc; 1.
KW Calcium; Kinase; Magnesium; Transferase.
FT DOMAIN 132 278
FT CONFLICT 378 378 V -> M (in Ref. 2). BAC34908.
FT CONFLICT 467 467 D -> Y (in Ref. 2). BAC34908.
SQ SEQUENCE 531 AA; 59811 MW; 31FECS34C348AA0A CRC64;

Query Match

Best Local Similarity 85.1%; Score 2428; DB 1; Length 531;

Matches 450; Conservative 33; Mismatches 46; Indels 0; Gaps 0;

QY 26 MGATGAAPLOSVLVWVKOORCAVSEPARALLRWRSPGCGAGAGDACSVPVSEITAV 85
DB 1 MGAMGAAPLESVLVWVKRRRCVSEPARALLRWRSPGCGAGDARSVLVSEITAV 60
QY 86 EETDVHGKHGSGKWKQKEKYAFVTHCVKARHRWKWAQVTFWCPEQLCHLWLQTLR 145
DB 61 EEKDDCEKASSGRHWKKNENPFATVHRVKVRHHRWKWARTVFWSADEQLCHLWLQTLR 120
QY 146 EMLEKLSRPHKHLVFINPFGKGQGRKRYERKVPAPLFTLASITTDIITVTEHANOAKETL 205
DB 121 GLLESLSRPHKHLVFINPFGKGQGRKRYERKVPAPLFTLASITTEIITEHANOAKETL 180
QY 206 YEINIDKVGIVCGVCGDMPSVHLGLIGRTORAGVDQNHPRVAVLVPSSLRIGIIPAGS 265
DB 181 YEINTSDYDGTVCVCGDMPSVHLGLIGRTVQAGIDPNHPRAVLVSTLRIGIIPAGS 240
QY 266 TDCVYSTVGTSDAETSAHHTVWGDLSAMDVSSVHHNSTLLRYSVSLGYGYGDIIDKDS 325
DB 241 TDCVYSTVGTNDATSAHHTVWGDLSAMDVSSVHHNSTLLRYSVSLGYGYGDIIDKDS 300
QY 326 EKKRWLGLARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQSKQ 385
DB 301 EKKRWMLGLVRYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQSKQ 360
QY 386 LEEQKALYGLAAEDVEEHWVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445
DB 361 LEEBKALYGLAAEDVEEHWVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
QY 446 ILIRKSRFNLRLRHTNQDQDFDTFFVBYRVKQFQFTSKHMEDESDSLKEGKKRF 505
DB 421 ILIRKSRFNLRLRHTNQDQDFDTFFVBYRVKQFQFTSKHMEDESDSLKEGKKRF 480
QY 506 GHICSSHPSCCTVSNSWNCGEVHLSPAIEVVRHCOLVRLFARGIEE 554
DB 481 GKICKDRPSCCTVSNSWNCGEVHLSPAIEVVRHCOLVRLFARGIEE 529
RESULT 3
Q6ZPK5 PRELIMINARY; PRT; 409 AA.
AC Q6ZPK5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE MKIAA1646 protein (Fragment).
GN Name=MKIAA1646;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryonic tail;
RX PubMed=14621295;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saka Y., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT III. the complete nucleotide sequences of 500 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:167-180(2003).
DR EMBL; AK129416; BAC98226.1; -;
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKc.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKc; 1.
DR SMART; SM00046; DAGKc; 1.
FT NON TER 1
SQ SEQUENCE 409 AA; 45697 MW; 810E815B37E26808 CRC64;

Query Match 62.9%; Score 1904; DB 2; Length 409;

Best Local Similarity 87.6%; Pred. No. 3.8e-147;

```
Matches 352; Conservative 26; Mismatches 24; Indels 0; Gaps 0;
QY 153 SRPKHLVFNPFPGKGGKRIYERKVAPLFTLASITTDIIIVTEHANOAKETLYEINDK 212
Db 6 SRPKHLVFNPFPGKGGKRIYERKVAPLFTLASITTEIIITEHANOAKETLYEINTDS 65
QY 213 YDGIIVCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVCYS 272
Db 66 YDGIIVCVGGDMFSEVLHGLIGRTQOSAGIDPNHPRAVLVPSTLRIGIIPAGSTDCVCYS 125
QY 273 TVGTSDAETSAIHIVVGDLSAMDVSSVHHNSTLLRYSVSLGIFYGFDIHKDSEKKRWLG 332
Db 126 TVGTNDAETSAHIIIGDSLAIQVSSVHHNSTLLRYSVSLGIFYGFDIHKDSEKKRWMG 185
QY 333 LARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRSQKQLEAEQK 392
Db 186 LVRVDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPRDNKPCRCAGCFVCRSQKQLEAEQK 245
QY 393 ALYGLEAAEDVEEQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDSDDLILIRKCS 452
Db 246 ALYGLEAAEDVEEQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDSDDLILIRKCS 305
QY 453 RFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGKKRFGHICSSH 512
Db 306 RFNLFRLIRHTNQDQDFTFVEVYRVKFKHFTSKHVEDRDNDSKEQKQKFGKICKDR 365
QY 513 PSCCCTVSNWNCDEVLHSPAIEVRVHCOLVRLFARGIEE 554
Db 366 PSCCSASRWNCDGEVHSPAIEVRVHCOLVRLFARGIEE 407

RESULT 4
Q6NX59 PRELIMINARY; PRT; 339 AA.
AC Q6NX59;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE CERK protein.
DE Name=CERK;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RA Strausberg R.;
Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC067255; AAH67255.1; -.
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR ProDom; PD005043; DAGKC; 1.
SQ SEQUENCE 339 AA; 37780 MW; A4C2ACDF2B6F3D0 CRC64;
Query Match 59.9%; Score 1811; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 1-2e-139;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 224 MFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVCYSTGTSDAETSA 283
Db 1 MFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVCYSTGTSDAETSA 60
QY 284 LHIIVVGDLSAMDVSSVHHNSTLLRYSVSLGIFYGFDIHKDSEKKRWLGARYDFSLGKLT 343
Db 61 LHIIVVGDLSAMDVSSVHHNSTLLRYSVSLGIFYGFDIHKDSEKKRWLGARYDFSLGKLT 120
QY 344 FLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRSQKQLEAEQKALYGLEAAEDV 403
Db 121 FLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRSQKQLEAEQKALYGLEAAEDV 180
QY 404 EEWQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDSDDLILIRKCSRFNLFRLIRH 463
Db 181 EEWQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDSDDLILIRKCSRFNLFRLIRH 240
QY 464 TNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGKKRFGHICSSHPSCCCTVSNSS 523
Db 241 TNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGKKRFGHICSSHPSCCCTVSNSS 300
QY 524 WNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 562
Db 301 WNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 339

RESULT 5
Q6GLV1 PRELIMINARY; PRT; 572 AA.
AC Q6GLV1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE MGC84197 protein.
DE Name=MGC84197;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
```

```

RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Brain;
RX  MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA  Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA  Richardson P.;
RT  "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT  initiative.";
RL  Dev. Dyn. 225:384-391(2002).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Brain;
RA  Klein S., Gerhard D.S.;
RA  Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RL  EMBL; BC074350; AAH74350.1; -
DR  GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR  GO; GO:0007205; P:protein kinase C activation; IEA.
DR  InterPro; IPR001206; DAGKC.
DR  Pfam; PF00781; DAGK_cat; 1.
DR  ProDom; PD005043; DAGKC; 1.
DR  SMART; SM00046; DAGKC; 1.
SQ  SEQUENCE 572 AA; 64899 MW; E3D22DB2C93B5B77 CRC64;

Query Match          56.6%; Score 1713; DB 2; Length 572;
Best Local Similarity 58.4%; Pred. No. 2.6e-131;
Matches 320; Conservative 81; Mismatches 117; Indels 30; Gaps 4;

QY  35 LQSVLWVKQORCAVLEPARALLRWGRSPGAG----APG----- 71
DB  12 LCSVLVSRQSEVTLDPGRSLLS-WRELPRHGRSRPGICLTRLVHQTGLVNRTR 70

QY  72 ---ADACSVPVSEIIAETDVHGHQSGKQMKPEYAFVHCVKRARRHMKWAQVT 128
DB  71 QPFGGVTPVTEIVSGEAEIDEKYNSMKWQHICKPRAFTVRVQVRKHWRCKEVT 130

QY  129 FWCPEQLCHLWQLTLEMLEKLTSPKHLVFNPFQKGGQGGKRIYERKVAFLTLASI 188
DB  131 FWCSEQLFYQWLQAFHDLLEQOQTHRPKNLLVYINPYGKKRGKQIYENKVAFLPSAAGI 190

QY  189 TTDIVTETHANAKETLYEINIDKVDGIVCVGGDMGFSEVLHGLIGRTQRSAGVDONHPR 248
DB  191 CADVITVEYANVARDHLVDANLEKVDGVVCGDGMGFSEVLHGLIVRMQKSDVDHNHPS 250

QY  249 AVLVPSSLRIGIIPAGSTDVCYCTVGTSDAETSALHIVGDSLAMDVSSVHHNSTLLRY 308
DB  251 AQLSRCNMRIGIIPAGSTDICYATVGINDPETSAHLIILGDCPLDVCVHYKRTFLKY 310

QY  309 SVSLILGYGYGDIIDKSEKRWGLIARVDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPRD 368
DB  311 SVSLILGYGYGDIIVLSEKKNRWLGPARYDVSGFKTLFTHHCYEGSVSQPAKWILGSPRD 370

QY  369 RKPCRCAGFCVCRSQKQLEEEKKALYLEAAEDVEQVCGKFLAINATNMSCACRRS 428
DB  371 QTTCTISGVCICQSSKQLDEQQTQACSEHREQDDWTTKGRFMALNAVMSACACPRS 430

QY  429 PRLGSPAHLGSGSLLILRKSRFNFLRLIRHTNQDQDFTFVEYRVKPFQFTSK 488
DB  431 PNLGSPAHLADGSADLLILVKRSLDLFLRLIRHTNSKQDQDFPFVEYRVKVFQFTPK 490

QY  489 HWEDED---SDLKEGKKRFGHICSHSPSCCCTVSNSSWNCNDCGEVLHSPAEVRVHCQIV 545
DB  491 HPEDENESTDGIILGKKNFQICTDHPSCGCGNHVNSIWNCDGETLDQTAIEMRVHCQLI 550

QY  546 RLFAEGIE 553
DB  551 KLFARGIE 558

RESULT 6
Q6GMF3
ID Q6GMF3 PRELIMINARY; PRT; 485 AA.
AC Q6GMF3;

```

```

DT  05-JUL-2004 (TrEMBLrel. 27, Created)
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE  MG81777 protein.
GN  Name=MG81777;
OS  Xenopus laevis (African clawed frog).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC  Xenopodinae; Xenopus.
OX  NCBI_TaxID=8355;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Embryo;
RX  MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA  Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA  Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchéz A.,
RA  Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA  Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA  Rodriguez A.C., Grinchwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA  Kryzyski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA  Jones S.J., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
RT  and mouse cDNA sequences.";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Embryo;
RX  MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA  Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA  Richardson P.;
RT  "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT  initiative.";
RL  Dev. Dyn. 225:384-391(2002).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Embryo;
RX  Klein S., Gerhard D.S.;
RL  Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC074110; AAH74110.1; -
DR  GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR  GO; GO:0007205; P:protein kinase C activation; IEA.
DR  InterPro; IPR001206; DAGKC.
DR  Pfam; PF00781; DAGK_cat; 1.
DR  ProDom; PD005043; DAGKC; 1.
DR  SMART; SM00046; DAGKC; 1.
SQ  SEQUENCE 485 AA; 54874 MW; E7C30C87AE52214F CRC64;

```

```

Query Match          49.4%; Score 1495.5; DB 2; Length 485;
Best Local Similarity 55.3%; Pred. No. 1.3e-113;
Matches 280; Conservative 82; Mismatches 93; Indels 51; Gaps 5;

QY  2 EAANGPAPLGVAPPAMWTSPPAEMGATGAAPLOSLVWVKQORCAVLEPARALLRW 61
DB  2 ESAHSPLP-----SVLCSVLVSRQSEVTLDPGRSLLS-WR 37

QY  62 SPGPGAG-----APG-----ADACSVPVSEIIAETDVHGHQ 95
DB  38 ELRPRRGKDRSRPGICMTVRFHQAGLVNGTHFPSSGGVTIPTEIVSGEAEIDEKY 97

QY  96 GSGKQKMKPEYAFVHCVKRARRHMKWAQVTFCPEQLCHLWQLTLEMLEKLTSPR 155
DB  98 NSMKQKQMKPEYAFVHCVKRARRHMKWAQVTFCPEQLCHLWQLTLEMLEKLTSPR 157

```

```
Qy 156 KHLVFINPFGKGKGIYERKVPAPLFTLASITTTIIVTEHANOAKETLYEINIDKYDG 215
Db 158 KYLLVYINPFGKGRKGIYETKVPAPLPSAGIACADVITVETVANHARDNLVEVNLKYDG 217
Qy 216 IVCVGGDGMFSEVLHGLIGRTORSAGUNDQNHPRAVLPSSIRIGIIPAGSTDCVCYSTVG 275
Db 218 VVCVGGDGIFFSEVLHGLGRVKGSDVDHNNPNAHLSQCNIIRIGIIPAGSTDCICYATVG 277
Qy 276 TSDAETSAHIVVGDSLAMDVSSVHNSLTLRYSVSLGIFYGYDIIKDSKKWGLGLAR 335
Db 278 INDEPETSALHILGDCQPLDVCVSHNKRTFLKYSVSLGIFYGYDVLKGTENRWLGPAP 337
Qy 336 YDFSLGKFTLHSHHCYEGVTSFLPAQHTVGSPRDRKPCRCAGCFVCRSQKQLEEBQKXALY 395
Db 338 YDVSCKFTLHSHHCYEGSVSQPAKWVLGSPRDQTPCTSGCYICRQSKQLEEBQKQTVF 397
Qy 396 GLE-AAEDVEWQVVCVGHFLAINATNMSCACRSPRGLSPAHLGDGSSDILIRKCSRPF 454
Db 398 GSEHRGQDDDMWTIKGRFMAINAVMSACAPRTPKGLSPAHLADGSAIDLVRKCSRL 457
Qy 455 NFLRLFIHTNQOQDFTFVEVYRV 480
Db 458 DFLRLHRIHTSNKOQDFPFVEVYRV 483
Qy 487 NFRFLRIHTNQOQDFTFVEVYRV 480
Db 489 DFLRLHRIHTSNKOQDFPFVEVYRV 483
RESULT 7
Q9VNA6 PRELIMINARY; PRT; 687 AA.
AC Q9VNA6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE CG16708-PA (Cg16708-pb).
CN ORFNames=CG16708.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
```

```
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; Smith H.O.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RL melanogaster euchromatic sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RL a genomic perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RL systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003603; AAF52040.2; -.
DR FlyBase; FBgn0037315; CG16708.
DR GO; GO:0048102; P:autophagic cell death; IEP.
DR GO; GO:0035071; P:salivary gland cell death; IEP.
DR InterPro; IPR001206; DAGK.
DR Pfam; PF00781; DAGK_cat; 1.
DR SMART; SM00046; DAGKc; 1.
SQ SEQUENCE 687 AA; 75690 MW; AAE081230A939412 CRC64;
Query Match 26.5%; Score 801.5; DB 2; Length 687;
Best Local Similarity 31.3%; Pred. No. 1.3e-56;
Matches 197; Conservative 97; Mismatches 206; Indels 129; Gaps 16;
Qy 36 QSVLWVQQRCAVSL-PPARALLRWRSPPGPGAGPCADACSVFVSEIIIAVEETDVHGK 93
Db 59 QQLVWERLQTKQSPQNGEAKPL-----PPDSAPQGGICSYGFSQSHVLHD--DVVSI 111
Qy 94 HQSGKQWQMEKP-----YAFTHCVKRRAR-----HRWKWQ 126
Db 112 RSGDTKASSLKPPSPGSGSSGCGDVAQKPTSQVLTINYAMRLSKSDTCNREWLRR 171
Qy 127 VTFWCPCEQLCHLWLTREMLEKLT---SPKHLVFINPFGKGKGIYERKVPAPL 183
Db 172 LTFNSDPYIVRQMDQLQIRLHSSSPTRMRVRLVFINPYGGRKAGAQTYERHVFIF 231
```


Qy	184	TLASITTDIIIVTEHANQAKEYINIKYDGI VCVGGDMPESEVLHGLIGTQRSAGVD	243
Dd	232	QLAGVADATCIIITQRANQVKDILLSHDLGVIDAVCCVGDDGTVAEINGLIFFROMRELGLD	291
Qy	244	QNHPRAVLVPSSLRIGIIPAGSTDCVCYSTGTSDAETSAHLI VVGDLSLAMDVSSVHHNS	303
Dd	292	EORPPYIPRP-ALPGVVIPAGSTDIAYSMHGTADVRTAA IHVILGOHRGLDVCSVNQG	350
Qy	304	TLLRYSVSLG YGVGDIIKQSKRWIGLARIDFSG LKTFLSHHCYECTVSFL-----	357
Dd	351	SLLRFCA SVLYGYLDVAAQSENRYWMPRRYEYGVKAF LNNGRYDAELRMLSEPDLL	410
Qy	358	---PAQHTVGSPRD-----RKPCRACGFVC-----PQSKOOLEEEOKKALYLE	398
Dd	411	LTTPLEDI PQSPDSVCSLGESVPSVCYANCORCSF ASSIQBORSSLFIOESKEA----	466
Qy	399	AAEDVE-----EWOVVCCKFLAINAT	419
Dd	467	RNQOVETEDSHLAASEALLRPRPRGNRLRPTGSI SSMRN LGNDQWKVVRGNFFMCGA	526
Qy	420	NMSCACRSR PGLSPAHLGDGSDLLIRKCSRNFN FLRIHTNQO--DQDFDTFFVEVY	478
Dd	527	NITCACARSPNGI SYSHLGDGCLDILLVKKTS LLNNVRFLNTAGRSGDINLPFEVY	586
Qy	479	RVKKFQFTSKHMEDEBDSLKEGKKRFGHIC-----SSHPSCCCTVSNWNCDGE	529
Dd	587	RTREFRFTFSAEEDYSLAGS-----COPITPPEMTAHSS---STEFSNNCDGE	635
Qy	530	VLHSPALIEVRVHCQLVRLFARGIENPKP	558
Dd	636	VWTDLDTIMRSHQCQLIEVMFRGPHSYSPK	664

```

RESULT 8
Q95S15 PRELIMINARY; PRT; 487 AA.
Q95S15;
AC Q95S15;
DT 01-DEC-2001 (TREMBlRel. 19, Created)
DT 01-DEC-2001 (TREMBlRel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBlRel. 26, Last annotation update)
DE HL01538p.
GN ORFNames=CG16708;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxid=7227;
[1]
RN
RC SEQUENCE FROM N.A.
RA STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacleb J., Parasvas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY051001; AAL28549.1; -.
DR FlyBase; FBgn0037315; CG16708.
DR GO; GO:0048102; P:autophagic cell death; IEP.
DR GO; GO:0035071; P:salivary gland cell death; IEP.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
DR SMART; SM00046; DAGKC; 1.
SQ SEQUENCE 487 AA; 53711 MW; AAE71EC40354BB07 CRC64;

```

Qy	214	DGIVCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRITGIIPAGSTDCVCYST	273
Db	62	DAVCCVGGDGTVAEVLINGLIFQRMRELGLDEQRPPYIPRP-ALPVGVIIPAGSTDTIAYSM	120
Qy	274	VGISDAETSAIHIVGVGDSLAMDVSSVHHNSTLIARSVSLGIVGYFFGDIIDKSEKKRWLGL	333
Db	121	HGTADVRTAAIHVTLGQHRGLDVCVSNQGSLRIFCAVSLVGYLGDVAAGSENTYRMWGP	180
Qy	334	ARYDFSGLTKFLSHHCYEGTVSFL-----PAQHTVGSPRD-----RKPCRACG	376
Db	181	RRVEYSGVKAFLNNRGYDAELRMLBEPDPLLTTPLEDIPQSPDSCVSLGESVPSVCYANC	240
Qy	377	FVC-----RQSKQLEEBQKALYGLEAAEDVE-----	404
Db	241	QRCSFASSIQEQRSSLFQIBESKEA-----ERNQOVETEDSHLAASEALLRPRPRLNR	296
Qy	405	-----EQVVQCGKFLAINATNWSACRRSPRGLSPAHLGDGSSDLILRL	449
Db	297	LPTGSISSMRNLGNDQWKVRGNFMFIGANITCACRSPNGISRYSHLGDGLDLILVK	356
Qy	450	KCSRFNRLRLIRHTNQO-QDQDFTFVEVYRVKKFQFTSKEMEDSDSLKEGKKRFGHI	508
Db	357	KTSLNNVRFLNLTAGRSGDIRNLFPFVEVYTRFRFRFTFSASEDYSLAGS-----	408
Qy	509	C-----SSHPSCCCTVSNSSWNCDEVLHSPAIERVVHVCQLVRLFARGIEENPKP	558
Db	409	CQPIPTPEEMTAHSS---STEFSSWNCDEGVVTDLIDITMRSHCOLIEVFMRGPHSYSKP	464
RESULT 9			
Q7PRA8			
ID O7PRA8		PRELIMINARY: PRT; 410 AA.	

```

178 KVALPFTLASITDDIIUTEHANOAKE--TLYEINIDKYDGIVCVCGDGMFSEVLHGLTGR 235
100 YAKPLFLAGVDINLIITORACQIYDIYTSKSIILLDNDYGLVCCGGDGTFAELFNGLVTR 159
41 KTNRWRVRVAVALHNSEPRVELWYTNRLSSDLRD-QNRPKHLLFLNYPYGGKKVALALYER 99
118 KRGRWAQWQVTFWCFESQLCHLMQIUKEMLEKITSRPAHLLVFINPFGAQQGAKIIR 17

```

```

Qy 236 TORSAGVDQNHPRAVLVPESSLRIGIIPAGSTDCVCYSTVGTSDAETSALHIVVGDSLAMD 295
Db 160 TMDGIDIKYP-AVLPRNIPIGVIPAGSTDTVACCLNGTTDIKTCTIIHILGOHSLD 218
Qy 296 VSSVHH-----NSTLLRYSVSLGIGYFYGDIIKDSEKRWGLGARY 336
Db 219 ISAVYSADAAKDEGASPGAGTGRPRPOLLLKLFASALSYGLGDIAYDSEKYRWMPKRY 278
Qy 337 DFSGLKTLFHHCHYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRSQKQLEEBEOKKALYG 396
Db 279 DYSGKFLNARGYNAEIIVHLDRRGQDPNDVGRCLDKARCKRKYGRDCCGERASY- 337
Qy 397 LEAEDVEEWOVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGLLILIRKCSRNF 456
Db 338 ---EDDDTEPLVVRGKFLMVSGANISCSERSPOGSPYCHLGDGLDLVLVRHTSMFNN 394
Qy 457 LRFLIRHTNQ 466
Db 395 LRLLLTWTSK 404

RESULT 10
Q6USK2 PRELIMINARY; PRT; 608 AA.
AC Q6USK2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ceramide kinase.
GN Name=CBRK;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22959668; PubMed=14563678; DOI=10.1101/gad.1140503;
RA Liang H., Yao N., Song J.T., Luo S., Lu H., Greenberg J.T.;
RT "Ceramide modulate programmed cell death in plants.";
RL Genes Dev. 17:2636-2641(2003).
DR EMBL; AY362552; AAQ62904.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR InterPro; IPR001205; P:protein kinase C activation; IEA.
DR Pfam; PF00781; DAGK_cat.1.
DR ProDom; PD005043; DAGKc; 1.
KW Kinase.
SQ SEQUENCE 608 AA; 68462 MW; 65BB9353692D9CA6 CRC64;

Query Match 21.4%; Score 648; DB 2; Length 608;
Best Local Similarity 31.6%; Pred. No. 4.2e-44;
Matches 182; Conservative 68; Mismatches 188; Indels 138; Gaps 19;

Qy 77 VPVSIIIAVEETDV-----HGKHQSGKQKMEKPYAFTVHCVRARRH--RWK 123
Db 71 IKFSIYAVEFVSYGLVHSPKGLRHAXEKRERLLNTQEMRYFTVHGFSQSPKELWN 130
Qy 124 WAQVTFWCPBEQCHLWLTQTLREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLF 183
Db 131 LAAFTFGHMDLQTCQSWMDQLNYSLIKVERPRNLLVFVHPKSGKNGSKVME-TVSKIF 189
Qy 184 TLASITTTDIIIVTEHANOAKETLYEI---NIDKYDGIVCVGDGMFSEVLHG-LIGRTO-- 237
Db 190 IRAKNTKVIIVTERAGHAFDWMASIQNKELHTYDGLIIVAGDGFNEILNGVILSRKVP 249
Qy 238 -----RSAG-----VDQ--NHPRAVLVPSL----- 256
Db 250 LPPSPDSFNSVQSRGSSVPEPGDEVHETDQKEHP---LLPDSVQVWMNFRVTNGSCE 306
Qy 257 -----RIGIIPAGSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSV--- 299
Db 307 GIEDPDHPFSRPRFGLIPAGSTDAIIVNCTTGARDPVTSAHLIILGRKFLDAMQVVRW 366

```

```

Qy 300 --HHNSTL---LRYSVSLGIGYFYGDIIKDSEKRWGLGARYDFSGLKTFLSHHCYEGTV 354
Db 367 KTASTSTIEPIRYAASFAGYGFYGDVISESEKYRWMPKRYDYVGTKIFLKHRSYEAEV 426
Qy 355 SFLPAQ-----HT-----VGSPPDRK-PCRAGCFVCRSQKQLEEBEOKKALYLE 398
Db 427 MFEAESENSKASLHTRSKTWPFRNTTTRSEKILCRANCKICNS-----KVGWNSASTTLN 481
Qy 399 AAEDEVEMQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGLLILIRKCSRNFRLR 458
Db 482 CPCKTTCWCTKGRFLSIGAAVMSNRNERAPDGLVVDHLSDDGFLHULLIKDCSRPKYLW 541
Qy 459 FLIRHTNQ-QDQDFTFEVYVRVKFQFTSKHMEDESDLKEGKGRKFHGHICSHHPSCCC 517
Db 542 HLTELAKRGGEPLNFEFVEYHKTRAFTTS-----FG----- 573
Qy 518 TVNSSNWNCDEVLHSPAIEVRVHVCQLVRLFARGIE 553
Db 574 --EESVWNLDEIPEAHQLSAQVLRGLIPLFASGPE 607

```

RESULT 11

```

Q6H6H1 PRELIMINARY; PRT; 700 AA.
AC Q6H6H1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative ceramide kinase.
GN Name=P0519E06.23; Synonyms=QJ1003.B06.39;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_taxid=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
RT clone:P0519E06.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX Sasaki T., Matsumoto T., Yamamoto K.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AP005006; BAD25678.1; -.
DR EMBL; AP004676; BAD25337.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR InterPro; IPR001206; DAGKc.
DR ProDom; PD005043; DAGKc; 1.
KW Kinase.
SQ SEQUENCE 700 AA; 77961 MW; FBB2C2A3DF31C75D CRC64;

Query Match 20.5%; Score 620.5; DB 2; Length 700;
Best Local Similarity 28.1%; Pred. No. 9e-42;
Matches 193; Conservative 87; Mismatches 197; Indels 209; Gaps 26;

Qy 18 WRTSPAAB---MGATGAAPLQSVLWVKQRCVSLP-ARALLRWFRSPFGAGAPCAD 73
Db 59 WLMKSSGEDRYLGQHDIEEVSSCW-----SSIMQPKLESKLKF-----SD 101
Qy 74 ACSVPVSIIIAVEE-----TDVHGKQSGKQKMEKPYAFTVHCVRARRHVKW--AQ 126
Db 102 VYVELLEVGVCPEFWNARATVQKIN-----TENR---FVIHTVPRKRPSPVQCE 153
Qy 127 VTFWCPBEQCHLWLTQTLREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFLA 186
Db 154 YIFGHKQDQCTKWVHEIKTCINKEQDRPKSLMVFVHPLCGKGRCKNWE-TVAPLPERA 212
Qy 187 SITTDIIIVTEHANOAKETLYEI---NIDKYDGIVCV----- 219

```

```
Db 213 KVKTKVIVTQRAGHAVDTLASLDKDKKFDGVIANTINACLSLFDIKHNYKMSARPE 272
QY 220 -----GGDGMFSEVL----- 229
Db 273 NTLSDYDQSAAGSHKSMILFYCFIINNKKQBEHRNNDLSNSELTGDDANAISGSSNTPDD 332
QY 230 HGLIGRTORSAGVD-----QNHPRVLV--PSS-LRIGIIPAGSTDC-----V 269
Db 333 HEPLLSSTRSTGLDSSSDSDEPCNGDQVPLVSPFNWFRIGIIPSGSTDAIVLSPVDV 392
QY 270 CYSTVGTSDAETSALHIVVGDSLAMDVSIV-----HNSLTLLRYSVSLIGYGFYGD1 321
Db 393 C-STTGERDPVTSALLIILGRISLDIAQVVRWKSSPSAEVSPTRYAASFAGYGFYGEV 451
QY 322 IKDSEKRWLGLARVDFSLGKTFSLHCHYEGTVSFLPAQHT-----V 363
Db 452 IRESEKYRMGPARYDFSGTWVFLKHSRYEAKVAFLENGWTHSLTASAEENNANGVQTLQY 511
QY 364 GSPDRKP-CRAGCFVCR---QSKQLEEEQKALYGLEAAEDVEMQVCGKFLAINAT 419
Db 512 HQNRHKTKICRNLCLCKGTSTSEQNSEDENPDS---SRTACETPKVWMSKGRFLUSVGA 568
QY 420 NNSCACRSPRGLSPAHLGDSSDLILIRKCSRFLRLRHTNQ-----QDQDFDTF 474
Db 569 VTSNERAPDGLVADAHLSDFLHLLLRDCP-----LPFYLWHLTQFTKGSDDLSPKPF 624
QY 475 VEVYRVKQFTSGKMEDESDSLKGGKKRFGHICSSHPSCCCTVSNSSWNCDBGVHLSP 534
Db 625 VEHHTQTAFTTSSHDE-----SVNNLDGELLQAC 654
QY 535 AIEVRVHCQLVPLFARGIE--ENPKP 558
Db 655 EVSVQAFRLVNLFSAGPEKMENTAP 680
```

RESULT 12

```
QSUZP6 ID Q6UZP6 PRELIMINARY; PRT; 532 AA.
AC Q6UZP6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ceramide kinase-like protein.
GN Name=CERK;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14681825;
RA Tuson M., Marfany G., Gonzalez-Duarte R.;
RT "Mutation of CERK, a novel human ceramide kinase gene, causes
RT autosomal recessive retinitis pigmentosa (RP26).";
RL Am. J. Hum. Genet. 74:118-138 (2004).
DR EMBL; AY357073; AAR13670.1; -.
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKc.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKc; 1.
KW Kinase.
SQ SEQUENCE 532 AA; 59602 MW; C73E590F7C25BED1 CRC64;
```

```
Query Match 19.6%; Score 594; DB 2; Length 532;
Best Local Similarity 28.4%; Pred. No. 9.1e-40;
Matches 159; Conservative 107; Mismatches 226; Indels 68; Gaps 14;
QY 9 PLGVAPPAPWRTSPAEMGATGAAP--LQSVLWVKQRCVAVSLPBARALLRWWRSPGFG 66
Db 23 PEAAVPPALLTSP---QQTAAAEAILLRGIFEGROSCDVLS-ERALR--WRPIQPE 76
QY 67 AGAPADACSVPSVEIIAVEET-DVHGKHQGGKQKMEKPYAFTVHCVKARRHRWKWA 125
```

```
Db 77 RPAGDSKYDLCKEEFIEKDI FSVKLKRRCSVQQRSGTLLGITLFTCLKKEQNKLN 136
QY 126 QVTPWCPEQLCHLWQLTREMLEKLTSRPKHLVFIINPFGKGQKGRKYERKVAPLFTL 185
Db 137 TLDLINLSEDCDIWFRQFKILAGFPNRPKSLKILLNPQSHKKEATQVYVEKVEPLKL 196
QY 186 ASIYTDIIVTHANQAKETLYEINIDKYDGVICVGGDGMFSEVLHGLIGRTORSAGVDON 245
Db 197 AGIKTDVTIMYEYGHALSLLKECELCQGFDPVVCVGGDGSASEVAHALLRQAQKAGMETD 256
QY 246 HPRVLVP--SSLRIGIIPAGSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSVHNS 303
Db 257 R---ILTPVRAQLPLGLIPAGSTNVLAHSLHGVPHVITATLHIINGHVQLVDVCTFSTAG 313
QY 304 TLLRYSVLLGYGFYGDIIKDSKKRWLG-LARYDFSLGKTFSLHCHYEGTVSFLPAQHT 362
Db 314 KLLRFGFSAM-FGFGRTLALAEKYRWMSPNQRDFAVVKALAKLAEDCEISFLPFNSS 372
QY 363 VGSFDRKPCRAGCFVCRQSKQQLBEEQKALYGLEAAEDVEMQVCGKFLAINATNWS 422
Db 373 -----DDVQERRAQGSPKSDCNDQMOMIQGQFLNVSMAIP 408
QY 423 CACRRSPRGLSPAHLGDSSDLILIRKCSRFLRLRHTNQDQDQDFTFVEYVRVKK 482
Db 409 CLCSVAPRGLAPNTRLNNGSMALIAARTSRPEFIKHLKRYASVKNQNFPPFVETYTEE 468
QY 483 FQFTSKH-----MEDESDSLKGGKKRFGHICSSHPSCCCTVSNSSWNCDBGVHLH-SP 534
Db 469 VKVHPNNTGGVPEEEDETASE-----NC-----FPMNVGDGLMEVAS 508
QY 535 AIEVRVHCQLVRLFARGIEE 554
Db 509 EVHRLHPRILSYGGSMEE 528
```

RESULT 13

```
Q9LU45 ID Q9LU45 PRELIMINARY; PRT; 533 AA.
AC Q9LU45;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similarity to unknown protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63 (2000).
DR EMBL; AB023044; BAA97392.1; -.
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKc.
DR ProDom; PD005043; DAGKc; 1.
SQ SEQUENCE 533 AA; 60153 MW; 1B851C7606B03E0E CRC64;
```

```
Query Match 19.3%; Score 593.5; DB 2; Length 533;
Best Local Similarity 33.5%; Pred. No. 6.6e-39;
Matches 158; Conservative 57; Mismatches 157; Indels 99; Gaps 16;
QY 77 VPVSEIIAVERTDV-----HGKHQGGKQKMEKPYAFTVHCVKARRH--RWK 123
Db 71 IKFSDIYAVEFVSVGLVHSPKGLGRHAKCEPRLNLTQEMYRFTVHGFSQSPKPCLN 130
```


This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 6, 2005, 07:01:07 ; Search time 5404.27 Seconds
(without alignments)
3958.373 Million cell updates/sec

Title: US-10-631-958-11
Perfect score: 3025
Sequence: 1 HEANGPAPLGVAPPAWRT.....QLVRLFARGIENPKPDHS 562

Scoring table: BLOSUM62
-Q= Xgapop 10.0 , Xgapext 0.5
-R= Ygapop 10.0 , Ygapext 0.5
-S= Fgapop 6.0 , Fgapext 7.0
-D= Delop 6.0 , Delext 7.0
Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10631958/runat_02092005_165816_4562/app_query.fasta_1.1941
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CNC 1_1 9235 @runat_02092005_165816_4562 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2478.5	81.9	1824	3 AK042077	AK042077 Mus muscu
2	2469.5	81.6	4248	3 AK052269	AK052269 Mus muscu
3	1518	50.2	1063	4 BM479389	BM479389 AGENCOURT
4	1367	45.2	1059	5 BQ054406	BQ054406 AGENCOURT
5	1321	43.7	1078	5 BQ063738	BQ063738 AGENCOURT
6	1299.5	43.0	797	7 CK000755	CK000755 AGENCOURT
7	1284	42.4	1047	5 BQ057191	BQ057191 AGENCOURT
8	1275	42.1	713	7 CK000525	CK000525 AGENCOURT
9	1260	41.7	820	6 CD655311	CD655311 AGENCOURT

10	1244	41.1	732	7 CF135528	CF135528 UI-HF-BNO
11	1205	39.8	758	6 CB246749	CB246749 UI-M-FIO-
12	1181	39.0	653	5 BX952302	BX952302 DRFp781L
13	1170	38.7	661	7 CN296312	CN296312 170005321
14	1107	36.6	1017	5 BQ879739	BQ879739 AGENCOURT
15	1059	35.0	812	7 CK603033	CK603033 AGENCOURT
16	1040	34.4	582	5 BP274786	BP274786 BP274786
17	1012	33.5	581	5 BP224560	BP224560 BP224560
18	1008	33.3	550	2 AW499858	AW499858 UI-HF-BNO
19	997	33.0	584	5 BP310011	BP310011 BP310011
20	986	32.6	779	5 BU265377	BU265377 603374309
21	982	32.5	584	5 BP309990	BP309990 BP309990
22	971	32.1	763	7 CK364327	CK364327 AGENCOURT
23	968	32.0	573	7 CF138275	CF138275 UI-HF-BNO
24	964	31.9	819	7 CN235868	CN235868 WLB086C10
25	959	31.7	550	7 CF138634	CF138634 UI-HF-BNO
26	943.5	31.2	759	5 BU264162	BU264162 603814452
27	942	31.1	541	7 CN296311	CN296311 170005321
28	925	30.6	725	5 BU214294	BU214294 603756937
29	924	30.5	607	6 CB581157	CB581157 AMGNNUC.N
30	916	30.3	576	6 CB608292	CB608292 AMGNNUC.N
31	914	30.2	653	6 BY754042	BY754042 BY754042
32	909	30.0	608	6 CB580936	CB580936 AMGNNUC.N
33	876	29.0	553	6 CA578972	CA578972 K0731A08-
34	874	28.9	662	2 BB638867	BB638867 BB638867
35	873	28.9	674	1 AJ739034	AJ739034 AJ739034
36	849	28.1	902	5 BU373554	BU373554 603811294
37	841	27.8	656	6 BY739924	BY739924 BY739924
38	838	27.7	545	6 CB611947	CB611947 AMGNNUC.N
39	828	27.4	488	5 BU430459	BU430459 UI-HF-BNO
40	813	26.9	540	6 CB612897	CB612897 AMGNNUC.N
41	811	26.8	474	2 AW503999	AW503999 UI-HF-BNO
42	806	26.6	638	1 AJ739022	AJ739022 AJ739022
43	805	26.6	449	2 AW503483	AW503483 UI-HF-BNO
44	797.5	26.4	759	5 BU290807	BU290807 604162346
45	786.5	26.0	680	2 BB235873	BB235873 BB235873

ALIGNMENTS

AK042077 1824 bp mRNA linear HTC 03-APR-2004
Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone: A630056D11 product: DAS9H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence.

ACCESSION AK042077 GI:26334912
VERSION AK042077.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE

Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL

REFERENCE

AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL

REFERENCE

AUTHORS

6 (bases 1 to 1824)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Igih,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurahara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Niehi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saichou,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission

TITLE

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES

source

Location/Qualifiers
1..1824
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:A630056D11"
/db_xref="taxon:10090"
/clone="A630056D11"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="3 days neonate"
85..1680

CDS

/note="unannoted protein product; DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTP[Q9UG55, evidence: FASTV, 86.5%ID, 100%length, match=945) putative"
/codon_start=1
/protein_id="BAC31157.1"
/db_xref="GI:26334913"
/translation="MGAMGAAEPLHSLVWVKRRRCVSLSEPARALLRWRRSPPEGPSA

ORIGIN

Alignment Scores:

Pred. No.: 1,09e-219 Length: 1824
Score: 2478.50 Matches: 470
Percent Similarity: 89.05% Conservative: 34
Best Local Similarity: 83.04% Mismatches: 54
Query Match: 81.93% Indels: 8
DB: 3 Gaps: 1

US-10-631-958-11 (1-562) x AK042077 (1-1824)

Qy 3 AlaAlaAsnGlyProAla-ProLeuGlyValArgAlaProProAlaTrpArgThrSerPr 22
Db 13 GCCGCTAACGGTCCCGCGCCCCCGGTGTCGCGCGCCCGCGCCCGCGCGGAGCGGCC 72
Qy 22 o--AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpVal 41
Db 73 GCTCTTGGCGGATGGGGGCAATGGGGGCGGGGCGGCGGTGCATCTCGTCTGGGTG 132
Qy 42 LysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArg 61
Db 133 AAACGGCGCGCTGTGCGCGTCAGCGTGGAGCCCGCGCGGCTCTGCTACGCTGTGGCGG 192
Qy 62 SerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGlu 81
Db 193 AGCCCGAGGCGCGGCCCTCCGCGCGGGTCCGATGCCCGCTCGGTACTGGTCTCGAG 252
Qy 82 IleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGln 101
Db 253 ATCATCGCTGTTGAGGAAAGAGACGCTGCAGAGAAACACGCGCTCAGTGGCGGATGCAT 312
Qy 102 LysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArg 121
Db 313 AAAATGAAAAATCGTTTCGCATTTCACAGTCACCGTCGTAAGCGAGTACGACATCACCGC 372
Qy 122 TrpLysTrpAlaGlnValPheTrpCysProGluGlnLeuCysHisLeuTrpLeu 141
Db 373 TGGAAAGTGGGACGGGTGACCTTCTGGAGCGCCGACGAGCAGCTGTGTACCTGTGGCTG 432
Qy 142 GlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPhe 161
Db 433 CAGACCTTCGTGGCTGTGGAGAGCTGACTTCAAGACCGAGACACTTGTGCTATTTC 492
Qy 162 IleAsnProPheGlyLysGlyGlnGlyLysArgIleTrpGluArgLysValAlaPro 181
Db 493 ATCAACCTTCGGAGGAAAGGTTCAGGCAAGCGCATCTATGAAAAAACAGTGGCGCCT 552
Qy 182 LeuPheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAla 201
Db 553 CTGTTTACCTTGGCTTCCATCCTACGAGATCATCAATTACAGAGCATGCCACAGAGGCC 612
Qy 202 LysGluThrLeuTrpGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGly 221
Db 613 AAGGAGACTTATACGAGATCAACACAGACAGCTATGATGCGCATCGTGTGCTAGGTGGG 672
Qy 222 AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly 241
Db 673 GACGGCATGTTTCAGCGAGGTGCTCATCGGGGTGATTGGGAGGACGACAGCGAGCGGTGT 732
Qy 242 ValAspGlnHisProArgAlaValIleuValProSerSerLeuArgIleGlyIle 261
Db 733 ATCAGCCCAATACCCCGAGCGCGTGTGGTGGCCAGTACCTCCAGATCGGCATCAT 792

PGDARSVLVSEIIAIVEEKDDCEKHASSGRWHKVENPFAFVHRVHRVHRHWRKWARV
TFWSAQOLCHLWLTQLGLLESITRSPKHLVFINPFGKGQCKRIYKTVAPLFTL
ASITTEIIITEHANOAKETUYENTSDYGVICVGGDMFSEVLHGVIGRTQOSAGID
PNHRAVLVPSTLRIGIIPAGSTDCVYSTVGTNDAAETSAHIIIGDSLALDVSSVHV
HNTLLRYSVLLGGYFGYDGLIKDSEKRWMLVRYDFSGLUKTFLSHYEGTSLFLPA
OHTVGSFRDNKPCRAGFCVCRSQKQLEEEKKALYGLNAEEMEEQVTCGKFLAIN
ATNMSCACPRSPGGLSPFAHLGSGSGLIILIRKSRNFRLFLRIHTNQDQDFTFV
EYVRVKFHFTSKHVEDNDNSKEQBKQKFGKICKDRPSCSASRSSMNCDEGVMS
PAIEVRVHCQLVRLFARGIEES"

Qy	262	ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr	281
Db	793	CCCCAGGGTCCACAGATTGTGTGTACTCAACAGTGGGCACAAACGACGACGACGAC	852
Qy	282	SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHis	301
Db	853	TCGGCTTTGCACATCATATTATTTGGGACCTCACTGGCAATAGACGTCTCTCTGTGCACCTAC	912
Qy	302	AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle	321
Db	913	CATAACACGCTGCTCGGTACTCGGTTTCTCTGCTGGCTACCGGTTTCTACGGGCACTTA	972
Qy	322	IleLysAspSerGluLysLeuValArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu	341
Db	973	ATCAAGGACAGTGAAGAAACCGGTGATGGCCCTCGCCGGTATGATTTCTCAGGGTTG	1032
Qy	342	LysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis	361
Db	1033	AAGACCTTCTCTCATCAGTACTATGAGGGACACTGTCTTCTCCAGCACGAC	1092
Qy	362	ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln	381
Db	1093	ACGGTGGGATCTCCACGGGACAAATAAACCTGCGGGCTGGGTCTCTGTGTCAGGCAG	1152
Qy	382	SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaGlu	401
Db	1153	AGCAAGCAACAGCTGGAAGAGAGAGAAAGCCCTGTATGGCTCGGAACGCCGAG	1212
Qy	402	AspValGluGluTTPGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet	421
Db	1213	GAATGGAGAGTGGCAAGTACATGTGGGAAGTTCCTGGCCATCATGCCACCAACATG	1272
Qy	422	SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly	441
Db	1273	TCCTGTGCTGTGCTCGGAGCCCTGGGGCCCTGTCCCCATTTCGCCCATCTGGAGATGG	1332
Qy	442	SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle	461
Db	1333	TCCTTCGACCTCATCTTATCCGGAAGTGTCCAGGTTCAACTTCCTCGAGATTCTCTATC	1392
Qy	462	ArgHisThrAnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys	481
Db	1393	CGGCACACGACACGAGGACCACTTCACCTTCACCTTCGTTGAAGTTATTCGAGTCAAG	1452
Qy	482	LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGly	501
Db	1453	AAATTCCTCATCTCACGTCCAGCACGTGGAAGACGAGGACAATGACTCGAAGGAACAAGAG	1512
Qy	502	LysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsn	521
Db	1513	AAGCAGAAGTTTGGGAAGATCTGCAAGGACACAGACCTCTTGCACTTGCTCAGCCCTCCAGA	1572
Qy	522	SerSerTrpAnCysAspGlyValLeuHisSerProAlaIleGluValArgValHis	541
Db	1573	AGCTCTGGAAGTGGGAGTCAATGACACAGCCCGCCATTCAGGTTCAGGTTCAC	1632
Qy	542	CysGlnLeuValArgLeuPheAlaArgGlyIleGluGlu-----AsnPro	556
Db	1633	TGCCAGCTGGTGGCCCTCTTCTCGGGGAATCGAGGAAGAGTCATAGCAAGAACCCCA	1692
Qy	557	LysProAspSer	560
Db	1693	AAGCCCGAGGAC	1704
RESULT 2			
AK052269			
LOCUS	AK052269	4248 bp	mRNA linear HTC 03-APR-2004
DEFINITION	Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched library, clone:D330016D08 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence.		
ACCESSION	AK052269		
VERSION	AK052269.1	GI:26342491	

KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi; Mus. Carninci, P. and Hayashizaki, Y.
AUTHORS	High-efficiency full-length cDNA cloning
TITLE	Meth. Enzymol. 303, 19-44 (1999)
JOURNAL	99279253
MEDLINE	10349636
PUBMED	10349636
REFERENCE	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishigaki, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
MEDLINE	
PUBMED	
REFERENCE	
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	
PUBMED	
REFERENCE	
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers
FEATURES	


```

Db      1748 AGCTCTGGAAGTGGGAGTGCAGAGCCCGGCAATGAGGTGAGGTCCAC 1807
      |||
Qy      542 CysGlnLeuValArgLeuPheAlaArgGlyIleGluGlu-----AsnPro 556
      |||
Db      1808 TGCCAGCTGGTGGCTCTTTGCTCGGGAATCGAGGAAGAGTCATAGCAAGAACCCCA 1867
      |||
Qy      557 LysProAspSer 560
      |||
Db      1868 AAGCCAGGAGC 1879

```

```

RESULT 3
LOCUS   BM479389
DEFINITION AGENCOURT_6418742 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5502281
5', mRNA sequence.
ACCESSION BM479389
VERSION   BM479389.1 GI:18528431
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE 1 (bases 1 to 1063)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12140 row: c column: 18
High quality sequence stop: 665.
Location/Qualifiers
1. .1063
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5502281"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

```

FEATURES

source

Alignment Scores:

```

Pred. No.: 1.62e-130 Length: 1063
Score: 1518.00 Matches: 301
Percent Similarity: 92.22% Conservative: 7
Best Local Similarity: 90.12% Mismatches: 19
Query Match: 50.18% Indels: 7
DB: 4 Gaps: 3

```

US-10-631-958-11 (1-562) x BM479389 (1-1063)

```

Qy      180 AlaProLeuPheThrLeuAlaSerIleThrAspIleValThrGluHisAlaAsn 199
      |||
Db      1 GCACCACTGTTACCTTACCTCCATCCACCTGACATCATCTGTTACTGAACTGCTAAT 60
      |||
Qy      200 GlnAlaLysGluThrLeuPyrGluIleAsnIleAspLysTyThrAspGlyIleValCysVal 219
      |||
Db      61 CAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCGCATCTGTTGTC 120
      |||
Qy      220 GlyGlyAspGlyMetPheSerGluValLeuHisGlyIleValArgThrGlnArgSer 239
      |||

```

```

Db      121 GCGGAGATGTTATGTTTCAGCGAGGTGCTGCACGGTCTGTATTGGAGACCGCAGAGGACC 180
      |||
Qy      240 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 259
      |||
Db      181 GCGGGGTGCGACCAAGAACACCCCGGGCTGTGCTGGTCCCAGTAGCTCCGGATTGGA 240
      |||
Qy      260 IleIleProAlaGlySerThrAspCysValCysTyThrValGlyThrSerAspAla 279
      |||
Db      241 ATCATTTCCCGAGGGTCAACGAGCTGCGTGTGTACTCCACGCTGGGACCGACGCGCA 300
      |||
Qy      280 GluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerVal 299
      |||
Db      301 GAACCTCGGCGCTGCATATCGTTGTGGGACCTGCTGGCCATGATGTGCTCTCAGTC 360
      |||
Qy      300 HisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGly 319
      |||
Db      361 CACCACAACAGCACACTCTTCCTGCTACCTCCGCTGCTGCTGGCTACGGCTTCTACGGG 420
      |||
Qy      320 AspIleIleLysAspSerGluLysIleArgTyrLeuGlyLeuAlaArgTyrAspPheSer 339
      |||
Db      421 GACATCATCAAGGACAGTGAGAAACGGTGGTGTGCTTGCAGATACGACTTTTCA 480
      |||
Qy      340 GlyLeuLysThrPheLeuSerHisHisCysTyThrGluGlyThrValSerPheLeuProAla 359
      |||
Db      481 GGTTAAGACCTTCCTCTCCACCACTGCTATGAGGAGGACAGTGTCTCTCTCCCTGCA 540
      |||
Qy      360 GlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCys 379
      |||
Db      541 CAACACAGGTGGGATCTCCAGGGGATAGGAAGCCCTCCCGGGCAGGATGCTTTGTTGC 600
      |||
Qy      380 ArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAla 399
      |||
Db      601 AGGCAAAACAAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCAGCTATGTGTTTGAAGCT 660
      |||
Qy      400 AlaGluAspValGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThr 419
      |||
Db      661 GCGGAGGAGCTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCNCATCATGCCACA 720
      |||
Qy      420 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 439
      |||
Db      721 AACATGTCTGTGCTGTGTCGCGGAGCCCCANGGGGCTCTCNCCCGCTGCCACATTNGA 780
      |||
Qy      440 AspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPhe 459
      |||
Db      781 GACGGGTCTTCTGACCTCATCTATCCGGAATGCTCCAGGNTCAATTTCTGAGATT 840
      |||
Qy      460 LeuIleAspGlnHisThrAsnGlnGlnAspGlnPheAspPheThrPhe---ValGluValTyr 478
      |||
Db      841 CTATCAGGCACACCAACCCAGCAGGACAGTTTGACTTCCCTTTTGGTTGAGTTTATC 900
      |||
Qy      479 ArgValLysLysPhe-GlnPheThrSerLysHisMetGluAsp-----GluAspSerAs 496
      |||
Db      901 CCGGTGAGGAAAAATCCAGTTTACCTCCAAAACCCCTTTGGGAGGATGAAGGCACAGCGA 960
      |||
Qy      496 pleuLysGluGlyGly-----LysLysArgPheGly 506
      |||
Db      961 CCTCAAGGGGGGGGGGGGGGAAAAAAGCCCTTTTGGG 1000
      |||
RESULT 4
LOCUS   BQ054406
DEFINITION AGENCOURT_6771237 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5803668
5', mRNA sequence.
ACCESSION BQ054406
VERSION   BQ054406.1 GI:19813746
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1059)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)

```

JOURNAL
COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Lou Staudt
 cDNA Library Preparation: Ruben Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LNCM2040 row: a column: 13
 High quality sequence stop: 601.
 Location/Qualifiers
 1..1059
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5803668"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 99"
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 1,92e-116 Length: 1059
 Score: 1367.00 Matches: 276
 Percent Similarity: 92.74% Conservative: 5
 Best Local Similarity: 91.03% Mismatches: 14
 Query Match: 45.19% Indels: 8
 DB: 5 Gaps: 4

US-10-631-958-11 (1-562) x BQ054406 (1-1059)

QY 158 LeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArg 177
 Db 15 TTACTGGTATTATCAACCGGTTTGAGGAAAAGGACAGGCAAGCGATATGAAAGA 74
 QY 178 LysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleValThrGluHis 197
 Db 75 AAAGTGGCACCACCTGTTCCACTTAGCCTCCATCACCACCTGATCATCTGTTACTGAACAT 134
 QY 198 AlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleVal 217
 Db 135 GCTATCAGGCCAAGGAGACTCTGTATGAGATTATACATAGAAATACAGCGGATCGTC 194
 QY 218 CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln 237
 Db 195 TGTGTCCGGGAGATGGTATGTTCCAGCAGGTGCTGCACGGTCTGATTGGGAGGACGCG 254
 QY 238 ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg 257
 Db 255 AGGAGCGCGGGGTCCAGCAGAACACACCCCGGGCTGTGCTGCTCCCGTAGCTCCGG 314
 QY 258 IleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSer 277
 Db 315 ATTGGAATCATTCCTCCGCGGGTCAACGAGCTGCGGTGTGTTCTCCACCGTGGGACACG 374
 QY 278 AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSer 297
 Db 375 GACGCAGAAACCTCGCGCTGCATATCGTTGTGGGGACTCGTGGCCATGATGTGTCC 434
 QY 298 SerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 317
 Db 435 TCAGTCCACCAACACAGCACACTCCTTCGCTACTCCGTGTCCTGCTGGGCTACGGCTTC 494

QY 318 TyrGlyAspIleIleLysAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAsp 337
 Db 495 TACGGGACATCATCAAGGACAGTGAGAAACGGTGGTGGTCTTGCAGATACGAC 554
 QY 338 PheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeu 357
 Db 555 TTTTCAGGTTAAAGACCTTCTCTCCACCACTGCTATGAACGACAGTGTCTCTCTC 614
 QY 358 ProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 377
 Db 615 CCTGCACAACACACGCGTGGGATCTCCAAGGATAGGAAGCCCTCCGGGCGAGGATGCTTT 674
 QY 378 ValCysArgGlnSerLysGlnLeuGluGluGlnLysLysLysAlaLeuTyrGlyLeu 397
 Db 675 GTTTGACGCAAGACAGCAGCAGCTGAGGAGGAGCAGAAAGCACTGTATGGGTG 734
 QY 398 GluAlaAlaGluAspValGluGluTyrGlnValValCys-GlyLysPheLeuAlaIleAs 417
 Db 735 GMAACTCGGAGGACGTGGAGGAATGGCAATGCTCTGTGTGGAAAGTTCTCTGGCCATTAA 794
 QY 417 nAlaThr-AsnMetSerCysAla---CysArgArgSerProArgGlyLeuSerProAlaA 436
 Db 795 TGCCACAAAACATGCCCTGTGCTTGTCCGCCCGACCCCTCCAGGGGCCCTCTCCCGCG 854
 QY 436 la-----HisLeuGly---AspGlySerSerAspLeuIleLeuIle-----ArgLysC 451
 Db 855 GCTGGCCCACTTGGGAAGACGGGTTCTTCCGACCTCTCTTCTTCTTCCTTCGGGAAT 914
 QY 451 ysSer 452
 Db 915 GCTCC 919

RESULT 5

BQ063738

LOCUS

DEFINITION

AGENCOURT 6873251 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925382

5', mRNA sequence.

ACCESSION

BQ063738

VERSION

BQ063738.1 GI:19891754

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1078)

NIH-MGC <http://mgs.nci.nih.gov/>.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Ruben Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LNCM2096 row: p column: 23

High quality sequence stop: 640.

Location/Qualifiers

1..1078

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5925382"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_99"

/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size

FEATURES

source

1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 3,82e-112 Length: 1078
Score: 1321.00 Matches: 260
Percent Similarity: 92.31% Conservative: 4
Best Local Similarity: 90.91% Mismatches: 15
Query Match: 43.67% Indels: 7
DB: 5 Gaps: 4

US-10-631-958-11 (1-562) x BQ063738 (1-1078)

QY 115 LysArgAlaArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGlu 134
DB 1 AGAGAGCAGCAGCGGACCGCTGGAAAGTGGGCGCAGGTGACTTTCTGGTGTCCAGAGGAG 60
QY 135 GlnLeuCysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArg 154
DB 61 CAGCTGTGTCACTTGTGGCTGCAGACCCCTCGCGAGATGCTCGAAGAGCTGACGTCCAGA 120
QY 155 ProLysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 174
DB 121 CCAAAGCATTTTACTGGTATTTATCAACCGCTTGGAGGAAAGGACAAAGCGGAGATA 180
QY 175 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleVal 194
DB 181 TATGAAGAAAGTGGCACCCTGTTCACTTAGCTTCATCACCACCTGACATCATCGTT 240
QY 195 ThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAsp 214
DB 241 ACTGAACATGCTAATCAGCCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGAC 300
QY 215 GlyIleValCysValGlyCysGlyMetPheSerGluValLeuHisGlyLeuIleGly 234
DB 301 GGCATCGTCTGTGTGGCGGAGATGATGTTTTCAGCGAGGTCTGCACGGTCTGATTGGG 360
QY 235 ArgThrGlnArgSerAlaGlyValAspGlnAanHisProArgAlaValLeuValProSer 254
DB 361 AGGACGACAGAGAGCGCGGGTGCACCAAGACCCCGGGCTGTGTGTCTCCAGT 420
QY 255 SerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTyrSerThrVal 274
DB 421 AGCCTCCGATTGGAATCATTCGCGAGGTCAACGGACTGCGTGTGTACTCCACCGTG 480
QY 275 GlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMet 294
DB 481 GGCACACGACGACGAGAAACCTTCGGCGTGCATATCGTTTGGGGACTCGCTGCCCATG 540
QY 295 AspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGly 314
DB 541 GATGTCTCTAGTCCACCAACACAGCACCTCTCTCGCTACTCCGTGTCCTGTCTGGGC 600
QY 315 TyrGlyPheTyrGlyAspIleLeuLysAspSerGluLysLysArgTrpLeuGlyLeuAla 334
DB 601 TACGGCTTCTACGGGACATCATCAAGACAGTGAAGAAACGGTGTGGTCTTCC 660
QY 335 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrVal 354
DB 661 AGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGTATGAAGGACAGGG 720
QY 355 SerPheLeuProAlaGlnHisThrValGlySer---ProArgAsp-ArgLysProCysAr 373
DB 721 TCTTCTCTCTCTGTCACACACACAGCGGTGGGATCTCCAGGGGAAATAAGAGCCCTGCC 780
QY 373 gala---GlyCys-----PheValCysArgGlnSerLysGlnGlnLeuGluGluGlu--- 389
DB 781 GGGCCAGGGATCCCTTGGTTTGGTTCAGCGCCAGGCAAGCCAGCCAGCTTGGNAGCGGG 840

QY 390 ----GlnLysLysAla 393
DB 841 AGCCCAAGAAAGGCC 856

RESULT 6

CK000755
LOCUS CK000755
DEFINITION AGENCOURT_16363467 NIH_MGC_221 Homo sapiens cDNA clone
IMAGE:30707875 5', mRNA sequence.
ACCESSION CK000755
VERSION CK000755.1 GI:38526789
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 797)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov

Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM1073 row: c column: 20
High quality sequence stop: 656.

FEATURES

Location/Qualifiers
1..797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:30707875"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_221"
/note="Organ: mixed; Vector: pYX-Asc; Site:1: EcoRI;
Site 2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. Average insert size
4-5Kb. Adaptors 5' (AATCGGACGAGG)3' and 5'
(CTCTGTCGG)3'. 3' linker sequence - GCGCGCGTGAAGCC T18.
Sequencing primers 3' end: T3 promoter primer 5'
(ATTAAACCTCACTAAAGGA)3', 5' end: T7 promoter primer 5'
(TAATAGACTCACTATAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Note: this is a NIH_MGC
Library"

ORIGIN

Alignment Scores:
Pred. No.: 2,42e-110 Length: 797
Score: 1299.50 Matches: 250
Percent Similarity: 96.54% Conservative: 1
Best Local Similarity: 96.15% Mismatches: 8
Query Match: 42.96% Indels: 1
DB: 7 Gaps: 1

US-10-631-958-11 (1-562) x CK000755 (1-797)

QY 177 ArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleValThrGlu 196
DB 9 CGGCACGAGGACCACTGTTCACTTCCATCCACTGACATCATCTGTTACTGAA 68

Qy 197 HisAlaAsnGlnAlaLysGluThrLeuTyrGluLeuAsnIleAspLysTyrAspGlyLe 216
Db |||||
Qy 69 CATGTAATCAGGCCAAGAGACTCTGTATGAGATTAAATAGACAAATACGCGCATC 128
Db |||||
Qy 217 ValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThr 236
Db |||||
Qy 129 GTCTGTGTGCGCGGAGATGATGTTTACAGGAGGTGTCACCGGTCTGATTTGGGAGACG 188
Db |||||
Qy 237 GlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeu 256
Db |||||
Qy 189 CAGAGAGCGCGGGTGCAGACAGAACCCCGGGCTGTGCTGCTCCCGATGAGCCTC 248
Db |||||
Qy 257 ArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThr 276
Db |||||
Qy 249 CGGATTGGAATCATTTCCGCGAGGTCAACGCACTGCGTGCTTACTCACCGTGGGCACC 308
Db |||||
Qy 277 SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal 296
Db |||||
Qy 309 AGCAGCGCAGAAACCTCGCGCTGCATATCGTTTGTGGGAGACTCGCTGGCCATGGATGTG 368
Db |||||
Qy 297 SerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGly 316
Db |||||
Qy 369 TCCTCAGTCCACCAACAGACACACTCTTCGCTACTCCGTCCTCGCTGGGCTAGCGC 428
Db |||||
Qy 317 PheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyr 336
Db |||||
Qy 429 TTCTACGGGGACATCATCAAGGACAGTGAAGAAGCGGTGTTGGGTCTTGGCCAGATAC 488
Db |||||
Qy 337 AspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPhe 356
Db |||||
Qy 489 GACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTTC 548
Db |||||
Qy 357 LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 376
Db |||||
Qy 549 CTCCTGCACAAACACACGGTGCGATCTCCAAGGATAGGAAGCCCTCGCGGGCAGGATGC 608
Db |||||
Qy 377 PheValCysArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrGly 396
Db |||||
Qy 609 TTTGTTTGCAGCAAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGGT 668
Db |||||
Qy 397 LeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIle 416
Db |||||
Qy 669 TTGGAAGCTGCGGAGGAGCTGGAGAGTGCNCAGTCTGCTGCGGAAATTTCTCGGGCNA 728
Db |||||
Qy 417 AsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAla 436
Db |||||
Qy 729 TCATGCACAAACATGCTCTGTGTTGTCGCGGAGCCCA---GGGGCTCTCCCGCTGCC 785
Db |||||

RESULT 7
BQ057191
LOCUS BQ057191
DEFINITION AGNCOURT_6769595 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5812382
5', mRNA sequence.
ACCESSION BQ057191
VERSION BQ057191.1 GI:19816531
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1047)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLC2062 row: 1 column: 15
High quality sequence stop: 535.
FEATURES
source 1..1047
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5812382"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GSCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN

Alignment Scores:
Pred. No.: 1.03e-108 Length: 1047
Score: 1284.00 Matches: 274
Percent Similarity: 91.67% Conservative: 12
Best Local Similarity: 87.82% Mismatches: 13
Query Match: 42.45% Indels: 13
DB: 5 Gaps: 2
US-10-631-958-11 (1-562) x BQ057191 (1-1047)
Qy 158 LeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArg 177
Db |||||
Qy 13 TTACTGTATTATTCAACCCGTTTGGAGGAAGAAGCAAGCAAGCGATATATGAAGA 72
Db |||||
Qy 178 LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHis 197
Db |||||
Qy 73 AAAGTGGCACCACTGTTTCACTTACCTTAGCCTCCATCACCACCTGACATCATCTGTTACTGAACAT 132
Db |||||
Qy 198 AlaAsnGlnAlaLysGluThrLeuTyrGluLeuAsnIleAspLysTyrAspGlyIleVal 217
Db |||||
Qy 133 GCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTC 192
Db |||||
Qy 218 CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln 237
Db |||||
Qy 193 TGTGTGCGGAGAGATGTAATGTTTCAGCAGGTGTCGACCGGTCTGATTTGGAGACGCGAG 252
Db |||||
Qy 238 ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg 257
Db |||||
Qy 253 AGGAGCGCGGGTTCGACCAAGAACCCCGGGCTGTGCTGCTCCCGTAGCTCCGG 312
Db |||||
Qy 258 IleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSer 277
Db |||||
Qy 313 ATTGGAATCATTTCCGCGAGGTCAACGCACTGCGGTGTGTACTCCACCGTGGGCACCAGC 372
Db |||||
Qy 278 AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSer 297
Db |||||
Qy 373 GACCCAGAAACCTCGCGCTGCATATGTTGTTGGGACTCGCTGGCCATGGATGTGTCC 432
Db |||||
Qy 298 SerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 317
Db |||||
Qy 433 TCAGTCCACCAACACAGCACACTCTCTCGGTACTCTCGGTGCTCCCTGCTGGGTACGGCTTC 492
Db |||||
Qy 318 TyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAsp 337
Db |||||
Qy 493 TACGGGACATCATCAAGGACAGTGAAGAAGCGGTGGTGGGTCTTGGCAGATACGAC 552
Db |||||
Qy 338 PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGlu-GlyThrValSerPheLe 357
Db |||||
Qy 553 TTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAGGGGAGACAGTGTCTCTCT 612
Db |||||
Qy 357 uProAlaGlnHisThrValGlySer-ProArgAspArgLysProCys-ArgAlaGlyCys 376
Db |||||

```

Db      613 CCTGCACACACACGCTGGGATCTCCCAAGGATAGGAAACCTGCCGGGACAGATGC 672
Qy      377 PheVal-CysArgGlnSerLysGlnLeuGluGlu--GlnLysLysAlaLeuTyr 395
Db      673 TTTGTTTTCGCGCAAGCAAGCAGCAGCTGGAGGAGGAGCCCAAGAAACCACTTGTAT 732
Qy      396 GlyLeuGluAlaAlaGluAspValGluGluTTPGlnValValCys-GlyLysPheLeuAl 415
Db      733 GTTTTGGAACTTGGCGAAGACGGGAGGAGTGGCAAACTTCTGGGGGGAAGTTTCTGGC 792
Qy      415 alleAsnAlaThrAsnMetSerCys-AlaCysArgArgSerProArgGlyLeuSer--P 434
Db      793 CATCATGCCCAACCTGCTCTGGTGTCTTGTGCGGGAACCCCGAGGGGCTTCTCCCC 852
Qy      434 roAlaAlaHisLeuGlyAsp---GlySerSerAspLeuLe-LeuLe-ArgLysCys-S 452
Db      853 CGGCTGCCCAATTTGGGGAAGAGGTCTCTTCTGAACCTCTCTCTCTCTCTCTCTCTCGT 912
Qy      452 erArgPheAsnPhe 456
Db      913 CCCGGTTCCATTTT 926

```

```

RESULT 8
CK000525
LOCUS AGENCOURT_16369000 NIH_MGC_221 mRNA linear EST 26-NOV-2003
DEFINITION IMAGE:30708597 5', mRNA sequence.
ACCESSION CK000525
VERSION CK000525.1 GI:38526559
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL 1 (bases 1 to 713)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM1075 row: a column: 22
High quality sequence stop: 689.
Location/Qualifiers
1..713
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30708597"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_221"
/note="Organ: mixed; Vector: pYX-Aec; Site1: EcoRI;
Site 2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Aec vector. Average insert size
4-5Kb. Adaptors 5'(AATTCGGACAGG)3' and 5'd
(CCTCGTCCG)3'. 3' Linker sequence - GCGGCGCTGAGGCC T18.
Sequencing primers 3'end: T3 promoter primer 5'd
(ATTAACCTCCTACTAAGGA)3'. 5' End: T7 promoter primer 5'd

```

FEATURES

```

source
1..713
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30708597"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_221"
/note="Organ: mixed; Vector: pYX-Aec; Site1: EcoRI;
Site 2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Aec vector. Average insert size
4-5Kb. Adaptors 5'(AATTCGGACAGG)3' and 5'd
(CCTCGTCCG)3'. 3' Linker sequence - GCGGCGCTGAGGCC T18.
Sequencing primers 3'end: T3 promoter primer 5'd
(ATTAACCTCCTACTAAGGA)3'. 5' End: T7 promoter primer 5'd

```

```

(TAATACGACTCACTATAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Note: this is a NIH_MGC
Library"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 3.92e-108 Length: 713
Score: 1275.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 42.15% Indels: 0
Gaps: 7
DB:
US-10-631-958-11 (1-562) x CK000525 (1-713)
Qy 302 AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 321
Db 9 AACAGCACACCTCTCGCTACTCTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATC 68
Qy 322 IleLysAspSerGluLysLysArgTyrLeuAlaArgTyrAspPheSerGlyLeu 341
Db 69 ATCAAGGACAGTGAAGAAACGGGTGGTCTTGCAGATACGACTTTTCAGTTTA 128
Qy 342 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 361
Db 129 AAGACCTTCTCTCCACCACTGTATGAAGGACAGTGTCTCTCTCCCTGCACACAC 188
Qy 362 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 381
Db 189 ACGGTGGGATCTCCAGGATAGGAGCCCTGCCGGGAGGATGCTTTGTTGACGGCAA 248
Qy 382 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaGlu 401
Db 249 AGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGACCTGTATGTTTGGAAAGCTCGGAG 308
Qy 402 AspValGluGluTTPGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet 421
Db 309 GACGTGGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCATGCCACAAACATG 368
Qy 422 SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 441
Db 369 TCCTGTGTTGTCGCGGAGCCCGAGGGCTCTCCCGGGCTGCCACTTGGGAGACGG 428
Qy 442 SerSerAspLeuLeuLeuLysArgLysCysSerArgPheAsnPheLeuArgPheLeuLe 461
Db 429 TCTTCTGACCTCATCTCCATCCGGAATGCTCCAGTTCAATTTTCTGAGATTCTCATC 488
Qy 462 ArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys 481
Db 489 AGGCACACCAACACGAGGACAGTGTGACTTCACTTTTGTGAAGTTTATCGCGTCAAG 548
Qy 482 LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGly 501
Db 549 AATTTCCAGTTTACGTCGAGCACATGAGAGATGAGGACAGGACCTCAAGAGGGGGGG 608
Qy 502 LysLysArgPheGlyHisLysCysSerSerHisProSerCysCysThrValSerAsn 521
Db 609 AAGAAGCGCTTTGGGCACATTTGCAGCAGCCCTCTCTGCTGTCGACCGTCTCCAAC 668
Qy 522 SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIle 536
Db 669 AGCTCTGGAACCTGCGAGGGGAGGTCTCTGCACAGTCTCTGCCATC 713

```

RESULT 9

```

LOCUS CD655311
DEFINITION AGENCOURT_14552675 NIA Human HI Embryonic Stem Cell cDNA Library
ACCESSION CD655311
VERSION CD655311.1 GI:31895467
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

CD655311 820 bp mRNA linear EST 18-JUN-2003
(AGNCOURT_14552675 NIA Human HI Embryonic Stem Cell cDNA Library
Long) Homo sapiens cDNA clone IMAGE:30426593 5', mRNA sequence.

```


Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 820)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: NDAM512 row: k column: 18
High quality sequence stop: 673.

FEATURES

1. Location/Qualifiers

1.20

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:10426593"

/tissue_type="Embryonic Stem cells"

/cell_line="WA01"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"

/notes="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, TRP1, TRP2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FUT3, SSEA-1, TRUB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 x 6 cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199] Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen: 5'-pGATAGTCTCATGCGCGCGCGCCCTTTTTTTTTTTT-3') from 3-4 µg of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loner-linker IL-SalI, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Tag polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-Sport6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Alignment Scores:		
Pred. No.:	1.22e-106	Length:
Score:	1360.00	Matches:
Percent Similarity:	93.13%	Conservative:
Best Local Similarity:	90.58%	Mismatches:
Query Match:	41.65%	Indels:
DB:	6	Gaps:
		820
		250
		7
		12
		5

US-10-631-958-11 (1-562) x CD655311 (1-820)

Qy	272	SerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSer	291
Db	12	TTCCACCGTGGGACACGACGACGAGAAACCTCGCGCTGCATATCGTTGTTGGGACATCG	71
Qy	292	LeuAlaMetAspValSerSerValHisHisAanSerThrLeuLeuArgTyrSerValSer	311
Db	72	CTGGCCATGGATGTGTCTCAGTCACCAACACAGCACACTCCTTCGTACTCCGTGTCC	131
Qy	312	LeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysLeuArgTyrLeu	331
Db	132	CTGTCTGGGCTACGGCTTCTACGGGACATCATCAGGACAGTGTGAGAAAGACGGTGTG	191
Qy	332	GlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGlu	351
Db	192	GGTCTTGGCAGATACGACTTTTCAGGTTTAGAGACCTTCCTCCACCACTGCTATGAA	251
Qy	352	GlyThrValSerPheLeuProIleGlnHisThrValGlySerProArgAspArgLysPro	371
Db	252	GGGACAGTGTCTTCTCCTCCGCAACACACGGTGGGATCTCCAAGGGATAGGAAGGCC	311
Qy	372	CysArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGluGluGluGlnLys	391
Db	312	TGCCGGGCGAGATGCTTTGTTGTCAGGCAACAGCAGCAGCTCGAGGAGGAGCAGAA	371
Qy	392	LysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGluTyrGlnValValCysGly	411
Db	372	AAAGCACTGTATGGTTTGAAGCTCGGAGGACGTGGAGGAGTGCACAGTCGTCTGTGG	431
Qy	412	LysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly	431
Db	432	AAAGTTTCTGGCCATCAATGCCAACACATGCTCTGTCTGTGTCGGGAGGCCAGGGGC	491
Qy	432	LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCys	451
Db	492	CTCTCCCGGCTGCCCATCTGGAGACGGGTCTTCTGACCTCATCTCATCCGGAATGC	551
Qy	452	SerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAsp	471
Db	552	TCCAGGTTCAAATTTTCTGAGATTTCTCATCAGGCACACCAACACGACGAGCAGTTGAC	611
Qy	472	PheThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGlu	491
Db	612	TTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTTCCAGTTTACGTCGAAGCACAATGAG	671
Qy	492	AspGluAspSerAspLeuLysGluGlyGlyLysLysArgPheGlyHisIleCysSerSer	511
Db	672	GATGAGGACAGGACCTCAAGGAGGNGGGAAGCGCTTGGGCACAT---TGCAGCAGC	728
Qy	512	HisProSerCysCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeu	531
Db	729	---ACCTCTGTCTGTGCAG---TCTCAGACTCTG---GACTGCGACGGGAGTCTG---	777
Qy	532	HisSerProAla---IleGluValArgValHisCysGlnLeuValArg	546
Db	778	-----CCAGCGTCATCATGTCAAGTCATGCACCTGTCTGATCTGCCAGA	819

RESULT 10

CF135528

LOCUS

DEFINITION

IM-HP-BN0-amf-g-10-0-UI.r1 NIH MGC 50 Homo sapiens cDNA clone

UTAGS:3090211 5', mRNA sequence.

CF135528	732 bp	mRNA	linear	EST 09-SEP-2003
LOCUS				
DEFINITION	UI-HF-BNO-amf-g-10-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone IMAGE:3090211 5', mRNA sequence.			

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Louis Staudt
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/humanfl.html
 Seq primer: pYX-5.

FEATURES
source

Location/Qualifiers
 1..732
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3090211"
 /tissue_type="lymph"
 /cell_type="germinal center B cells"
 /cell_line="MG85"
 /lab_host="DH10B (LTI)"
 /clone_lib="NIH MGC 50"
 /note="Vector: pT7T3-Pac; Site 1: NotI; Site 2: Eco RI;
 Constructed from size fractionated cytoplasmic mRNA
 (3.5-4.4kb). Directionally cloned. Cells provided by
 Louis M. Staudt, Ph.D. Library preparation by Maria de
 Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN

Alignment Scores:
 Pred. No.: 3,16e-105 Length: 732
 Score: 1244.00 Matches: 231
 Percent Similarity: 99.57% Conservative: 0
 Best Local Similarity: 99.57% Mismatches: 0
 Query Match: 41.12% Indels: 1
 DB: 7 Gaps: 0

US-10-631-958-11 (1-562) x CF135528 (1-732)

QY 332 GlyLeuAlaArgTyxAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrglu 351
 |||||
 Db 9 GGTCCTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGNA 68
 QY 352 GlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 371
 |||||
 Db 69 GGGACAGTGTCTCTCTCCCTCACACACACGCGTGGGATCTCCAGGGATAGGAAGGCC 128
 QY 372 CysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLys 391
 |||||
 Db 129 TCCCGGGCAGAGTCTTTGTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGGAGGAG 188
 QY 392 LysAlaLeuTyrglyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGly 411
 |||||
 Db 189 AAAGCACTGTATGGTTTGGAACTCGGAGGAGCTGGAGGATGGCAGTCTCTGTGGG 248
 QY 412 LysPheLeuAlaAlaAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly 431
 |||||
 Db 249 AAGTTTCTGGCCATCAATGCCAACACATGCTCTGTGCTGTGCGCGGAGCCCGAGGCGC 308
 QY 432 LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeuLeuLeuLeuLysCys 451
 |||||
 Db 309 CTCCTCCCGGCTGCCACTTGGAGACGGGTCTTCGACCTCATCTCATCCGGAATGC 368
 QY 452 SerArgPheAsnPheLeuArgPheLeuLeuArgHisThrAsnGlnGlnAspGlnPheAsp 471
 |||||
 Db 369 TCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACACACACAGCAGGACCATGTTGAC 428

QY 472 PheThrPheValGluValTyxArgValLysLysPheGlnPheThrSerLysHisMetGlu 491
 |||||
 Db 429 TTCACCTTTGTGAAGTTTATCGCGCTCAAGAAATTCAGTTTACGTGGAAGCACATGGAG 488
 QY 492 AspGluAspSerAspLeuLys-GluGlyGlyLysLysArgPheGlyHisLeuLysCysSerSe 511
 |||||
 Db 489 GATGAGGACAGCGACCTCAGGGAGGGGGGAGGAGGCGCTTGGGCACATTTGGCAGCAG 548
 QY 511 HisProSerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluVal 531
 |||||
 Db 549 CCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 608
 QY 531 HisSerProAlaLeuGluValAlaValHisCysGlnLeuValArgLeuPheAlaArgG 551
 |||||
 Db 609 GCACAGCCCTGTCATCGAGGTGAGAGTCCAGTCCAGCAGCTGGTGTGACTCTTTTGACAGG 668
 QY 551 yileGluGluAsnProLysProAspSerHisSer 562
 |||||
 Db 669 GATTGAAGAGATCCGAAGCCAGACTCACACAGC 702

RESULT 11

CB246749

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB246749 758 bp mRNA linear EST 09-JUL-2003
 UI-M-F10-cdx-b-10-0-UI.r1 NIH BMAP_F10 Mus musculus cDNA clone
 IMAGE:6835595 5', mRNA sequence.

CB246749 GI:28368393
 EST.
 Mus musculus (house mouse)

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

FEATURES
source

Location/Qualifiers

1..758

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6835595"

/tissue_type="whole brain"

/dev_stages="embryo 12.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP_F10"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is CAGCCACGAC. This library was created for the University

of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1 (bases 1 to 653)	
AUTHORS		Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.	
TITLE		EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)	
JOURNAL		Unpublished (2003)	
COMMENT		Contact: MIPS	
		MIPS	
		Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany	
		This is the 5' sequence of the clone insert	
		Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.	
		No sl sequence available.	
		This clone (DKFZp781l11183) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.	
FEATURES		Location/Qualifiers	
source		1..653	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="DKFZp781L11183"	
		/dev_stage="adult"	
		/lab_host="DH10B"	
		/clone_lib="781 (synonym: hlcc4)"	
		/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"	
ORIGIN			
Alignment Scores:			
Pred. No.:		1.98e-99	
Score:		Length: 653	
		1181.00	
Percent Similarity:		Matches: 217	
Best Local Similarity:		Conservative: 0	
Query Match:		Mismatches: 0	
		Indels: 0	
DB:		Gaps: 0	
US-10-631-958-11 (1-562) x BX952302 (1-653)			
Qy	322	IleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu	341
Db	2	ATCAGGACAGTGAGAAGAAACGGTGGTGGTCTGCCAGATACGATTTTCAGGTTTA	61
Qy	342	LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis	361
Db	62	AAGACCTTCTCTCCACCACTGCTATGAAGGGACAGTGTCTCTCTCCCTGCACAAAC	121
Qy	362	ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln	381
Db	122	ACGTGGGATCTCCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTGAGGCA	181
Qy	382	SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu	401
Db	182	AGCAAGACAGCAGCTGGAGGAGGACAGAAAGACACTGTATGTTGGTGGAGCTCGGAG	241
Qy	402	AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet	421
Db	242	GACGTGGAGGATGGCAAGTCTGTGGGAAGTTTCTGCCATCAATGCCACAAACATG	301
Qy	422	SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly	441
Db	302	TCCTGTGCTGTGTCGGGAGGCCCGGAGCCCTCTCCCGGCTGCCACTTGGGAGACGGG	361
Qy	442	SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle	461
Db	362	TCTTCTGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATC	421
Qy	462	ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys	481


```
Qy 61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSer 80
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 CGGAGCCCGGAGCCCGGCGCCCTCGCGGCCGATCCCGATGCTACTCGGTACCAAGTGTC 179
Qy 81 GluIleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 GAGATCATCGCGTTGAGGAAAGAGTCCACCGAAGAACAGTCTCCATGGCCGATGG 239
Qy 101 GlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHis 120
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
240 CATAGATGGAAATCCGTTTGATTCACAGTCCACTGCTGTGAGCGAGCTCGACACCAC 299
Qy 121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrp 140
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 CGCTGGAAGTGGCCACCGGTGACCTTCTGGAGCGCCGACGAGCAGCTGTCAACTGTGG 359
Qy 141 LeuGlnThrLeuArgGluMetLeuLysLeuThrSerArgProLysHisLeuLeuVal 160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
360 CTGCAGACCTCCGAGAGCTGCTGGAGAACCTGACTTCAAGACCGAGCATTTGCTGGTA 419
Qy 161 PheIleAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArgLysValAla 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
420 TTTATCAACCTTTCGGAGGGAAGGCCAGGCGACACATCTATGAAAAAAGTGCA 479
Qy 181 ProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGln 200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
480 CCTTTGTTCACTTGCTTCCATCACCACCTGAGATCATCATTTACTGAGCATGCCAATCAA 539
Qy 201 AlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 220
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
540 GCCAGGAGACTTTATACAGATCAACACAGACAGCTATGACGGCATCGTGTGTGTGT 599
Qy 221 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
600 GGGGACGCATGTTTCAGCGAGTGCTGCACGGGTAAATGGGAGACCGACAGCGCT 659
Qy 241 GlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 260
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
660 GGTGTTGACCAATACCCCGAGCCGCTGCTGTGTGCCAGTACCCT--AGGATCGGCATC 717
Qy 261 IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu 280
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
718 ATCCGGCAGGGTCACAGA---TGTGTGTC---TCTCAGGTGGCACAACG-----AGC 765
Qy 281 ThrSerAlaLeuHisIleValValGly 289
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
766 AGACGTCGCTTACCATCATATGGGA 792
```

Search completed: September 6, 2005, 20:30:03
Job time : 5431.27 secs

This Page Blank (uspto)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	567.5	18.8	549	2	T33517	hypothetical prote	
2	417.5	13.8	1240	2	T05162	hypothetical prote	
3	291.5	9.6	458	2	T38776	hypothetical prote	
4	248.5	8.2	687	2	S51398	hypothetical prote	
5	245.5	8.1	473	2	T19707	hypothetical prote	
6	210.5	7.0	624	2	S67059	hypothetical prote	
7	189.5	6.3	310	2	AG1665	hypothetical prote	
8	176	5.8	310	2	A11293	hypothetical prote	
9	147	4.9	303	2	P67935	conserved hypot	
10	136.5	4.5	306	2	AH1769	conserved hypot	
11	136	4.5	732	2	T16422	hypothetical prote	
12	130.5	4.3	309	2	AH1528	conserved hypot	
13	129	4.3	295	2	A83894	hypothetical prote	
14	126.5	4.2	306	2	A81394	conserved hypot	
15	125.5	4.1	295	2	D83734	hypothetical prote	
16	125.5	4.1	309	2	AF1171	conserved hypot	
17	122.5	4.0	433	2	S75948	hypothetical prote	
18	115.5	3.8	294	2	G95120	conserved hypot	
19	114	3.8	364	2	F84898	hypothetical prote	
20	112	3.7	345	2	E69578	involved in polyke	
21	111.5	3.7	311	2	C97990	conserved hypot	
22	105	3.5	333	2	F71006	hypothetical prote	
23	103.5	3.4	1028	2	A96719	hypothetical prote	
24	102.5	3.4	315	2	A89978	conserved hypot	
25	102	3.4	297	2	F95595	multidrug resistanc	
26	102	3.4	343	2	JC7183	cathelipin Q (EC 3.	
27	100	3.3	650	1	JC1450	fibroblast growth	
28	99	3.3	1555	2	T18688	hypothetical prote	
29	99	3.3	1973	2	G89608	protein B0272.5 [i	

Cy 523 SWNCDGEVLHSP-----AIEVRVHCQVLRFLAR 550
 |||||:-|-:::||-|-
Dd 496 VWNLDGSEILEQPDKDEPLHLKLPQLISFFGR 526
 :

RESULT 2

T05162
hypothetical protein F18E5.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05162
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Meuwissen, J.E.M.
submitted to the Protein Sequence Database, August 1998
A:Reference number: Z15400
A:Accession: T05162
A:Molecule type: DNA
A:Residues: 1-1240 <BE>
A:Cross-references: UNIPROT.O65419; EMBL.AL022603
A:Experimental source: cultivar Columbia; BAC clone F18E5
C:Genetics:
A:Map position: 4
A:Introns: 44/3; 117/3; 178/3; 206/1; 237/2; 375/1; 416/1; 449/3; 474/1; 503/1; 529/1; 532/1
A>Note: F18E5.160

Query Match 13.8%; Score 417.5; DB 2; Length 1240;
Best Local Similarity 28.9%; Pred. No. 1.3e+25;
Matches 129; Conservative 75; Mismatches 162; Indels 81; Gaps 16;

Cy 140 WLQTLEWLEKLTSRPKHLLVFNPFGKGQGKKRIYERKVAPLTLASITDTDIIVTEHAN 199
Dd 363 WCYKRLRYDLSL-GRPKRLLLVFNPFPGGKSAREIFVKVEKPLFEDADVQEIQTKEYQL 421
Cy 200 QAKETLYEINIDKYDGIVCVGGGMPESEVLHGILGRTO-QSAGVDQNHPRAVLVPSSLRI 258
Dd 422 HAKEFEVKSMDSKYDGIVCVGSGDIGLEVNGLLERADWRNA-----LKLPPI 468
Cy 259 GIIPAGSTDCVCYSTGTGSTD----AETSALHIVVGDSDLAMDVS SVHNHTLLRSYSLLG 314
Dd 469 GWFPAGTGNMIGSKSLDDTVGLGCCANSATSIIRGHKRSVDVTATIAQNTKF-FSVLMLA 527
Cy 315 YGFYGDI IKOSEKRWLGARYDF-----SGLKTFLSHHYEGTVSF 356
Dd 528 WGIIADI IDISEKFRFMWSARI DYPVCLVDKFDNYCIAVVKLLAQRIICLLRRYNCRILF 587
Cy 357 LPAQ--HTVGSPRDRKPCRCAGCFYCROSQQOLEEQKALYLEAAEDVEEMQVCGKEL 414
Dd 588 LPAFGFGYGPAS-----CSLYQBPHVSDKEVGOGPETKFEDLEWMKMGPfv 637
Cy 415 AINATNMSCARRSRPrg-----LSPAHLGDGSSDLILIRKC SRFNFLRI RHTNQODQ 469
Dd 638 TIWLHV-----PWGSENTLTAPA KPSDGYLDLVLKNCPKLVLLS-LMRQTSSGTH 689
Cy 470 FDFTFV-----EVYRVKKFOPT SKHMDEDSL KE GKGRFGHICSSHPSCCCTCNSSW 524
Dd 690 VESPYI VI VKLTVEKVAFL EPALGVDPD--KEGIIDS DGVELAR-----GKRTY 739
Cy 525 NCDEVLHS-PAIEVRVHCQVLRFLAR 550
Dd 740 KCDAQALMSYDKLQVTDQE--RLFCCK 764

RESULT 3

T38776
hypothetical protein SPAC4A8.07c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38776
R:Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21751
A:Accession: T38776
A>Status: Preliminary; translated from GB/EMLB/DDBB

[illegible]

Db 184 GK--LFQGEI-----MLFLVLTNSVGFGEK-----LAPDSSLNDGFMFLMIK 225
QY 450 KCSRNFNLFRLI-----RHTNQDQDFTFVEYRVK 481
Db 226 KANLAEFIRVATMALRGEHNDQ---HIIYTKANRVK 259
RESULT 10
AH1769
conserved hypothetical protein lin2702 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH1769
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1769
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-306 <GLA>
A:Cross-references: UNIPROT:Q92T76; GB:AL592022; PIDN:CAC97928.1; PID:g16415238; GSPDB:G
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2702

Query Match 4.5%; Score 136.5; DB 2; Length 306;
Best Local Similarity 19.6%; Pred. No. 0.0021;
Matches 62; Conservative 55; Mismatches 91; Indels 109; Gaps 14;
QY 156 KHLVFINPFGKGQKRI-----YERKVAPLFTLASITTDIIIVTEHANQAKE 203
Db 3 KAMIIYNPAGKNKFKLLPDAEKILTNADFEVLVP-STPAPKSTTLI-----AKQAAE 57
QY 204 TLYEINIDKYGIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPVSSLRIGIIPA 263
Db 58 AGYEV-----VIAAGDGTNVVNGLMQVEKRP-----KLGIPLV 93
QY 264 GSTDCVCVTSVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYS-----VSLLOYGYG 319
Db 94 GTTNDYARALNPAKDPLEALQIIAQETIRVDIGKANETEFINNAAGGRITEITYA--- 150
QY 320 DIIKDSEKKRWGLARYDFSGLTKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVC 379
Db 151 --VKESMSKW-GRLAYLFSGL-----TVLP----- 173
QY 380 RQSKQOLBEEQKALYGLEAABDEWQVVCVKFLAINATN-----MSCACRRSPRGLSPA 435
Db 174 -----KLSPVNVVEIYN---BEIFKGEIL---LFFVNKNTSVGGMETLC-----PP 213
QY 436 AHLGDGSSDLILIRKCS 452
Db 214 AOLNSGMFELLILKKVS 230

RESULT 11
T16422
hypothetical protein F52C9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16422
R:Pavello, T.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F52C9.
A:Reference number: Z18511
A:Accession: T16422
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-732 <PAV>
A:Cross-references: EMBL:U39850; NID:g1055052; PID:g1055058; PIDN:AAA81060.1; CESP:F52C9
C:Genetics:
A:Gene: CESP:F52C9.3
A:Introns: 63/1; 106/3; 148/2; 303/1; 347/1; 408/2; 650/2
Query Match 4.5%; Score 136; DB 2; Length 732;
Best Local Similarity 19.9%; Pred. No. 0.0075;
Matches 76; Conservative 56; Mismatches 128; Indels 122; Gaps 18;
QY 153 SRPKHLVFINPFGKGQKRIYERKVAPLFTLASITTDIIIVTEHANQAKETLYEINIDK 212
Db 65 TRPKRVFLVNVNVEGNSRCFPQFNKNALPLFLHLAGVQVDVVKADNQAOLEALAGAVDTQE 124
QY 213 YDGIYVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPVSSLRIGIIPAG----- 264
Db 125 ADILVVGDDGTIGTVTGTIFRNREK-----AQLPVGFYPGGYDNLWLKR 169
QY 265 -----STDCV---CYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLL--RYSVSLL 313
Db 170 MLPSVFNSDDVRHACETAMAVIEDQKSVY-----AFELTT--EGSTLAPEYGLGDV 220
QY 314 GYGFVGDIIKDSEKKRW-LGLAR---YDFSGLK---TFLSHHCYEGTVSFLPAQHTVG 364
Db 221 SAGWFRQ-IEDTRKKFWFSNAKRWAYFWEMLRGPAPIECHVEYET----- 268
QY 365 SPRDRKPCRAGCFVCROSQKQOLEBEQKALYGLEAABDEWQVVCVKFLAINATNMSCA 424
Db 269 -----C-AGCEKCR-PKPIIEAPQWR-----WWHVLTG----- 294
QY 425 CRRSPRGLSPAHLGDGSSDL--TLIRKCSRPNFL-----RFLIRHTNQDQDFTFVEV 477
Db 295 -----TPKYXNDGQKDYTGIIINEKCGKEHLDTHGAEFLIENEQMSD----- 337
QY 478 YRVKKFOFTSKHMEDESDLKE 499
Db 338 YSQIRFMGDPYMPSEEFWNE 359
RESULT 12
AH1528
conserved hypothetical protein lin0768 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH1528
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1528
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-309 <GLA>
A:Cross-references: UNIPROT:Q92DPS; GB:AL592022; PIDN:CAC96000.1; PID:g16413219; GSPDB:G
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0768
Query Match 4.3%; Score 130.5; DB 2; Length 309;
Best Local Similarity 22.3%; Pred. No. 0.0065;
Matches 42; Conservative 33; Mismatches 88; Indels 25; Gaps 5;
QY 159 LVFINPFGKGQKRIYERKVAPLFTLASITTDIIIVTEHANQAKETLYEINIDKVDGIVC 218
Db 5 LLIVNPSGSEKGR-TYQGTKEEVLKKRYDEVEVLTEKAGDATEFASWASEQGFDAVIA 63
QY 219 VGGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPVSSLRIGIIPAGSDVCYCVSTGVS 278
Db 64 MGGDGTNETINGL-----AIHEKRP-----DFGFIPLGTIVNLSRVGIFL 106

Db	162	YIKLEK-----LQI-----SPTDVR-----	178
Qy	397	LEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNF	456
Db	179	IBYDGKLFEGEIM--MFLVSN-TNSVGGFER----LAPNASLRDGMFDFIIVKKTSPPEF	231
Qy	457	LRFL-----IRHTNQDQDFTFEVEYRVK	481
Db	232	LHLAGLALRGEHIKPK-----VLYVQANRIK	258

Search completed: September 3, 2005, 04:38:45
Job time : 25.8463 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 5, 2005, 19:28:44 ; Search time 317.875 Seconds
(without alignments)
2892.921 Million cell updates/sec

Title: US-10-631-958-11
Perfect score: 3025
Sequence: 1 HEANGPAPLGVAPPWRT.....QLVRLPARGIENPKPDHS 562

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO pool/US10631958/runat 02092005.165811.4432/app query.fasta_1.1941
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0 1 -LOOPTCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=prco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CGN 1.1 397 @runat 02092005.165811.4432 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6CTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2931.5	96.9	4432	US-09-774-528-148	Sequence 148, App
2	813	26.9	2064	US-09-270-767-14306	Sequence 14306, A
3	585.5	19.4	1084	US-09-270-767-15155	Sequence 15155, A
4	302.5	10.0	901	US-09-270-767-30448	Sequence 30448, A
5	288.5	9.5	1857	US-09-970-516-3	Sequence 3, Appli
6	288.5	9.5	2380	US-09-817-676A-13	Sequence 13, Appl
7	280.5	9.3	1783	US-09-949-016-1155	Sequence 1155, Ap
8	269	8.9	2698	US-09-817-676A-11	Sequence 11, Appl
9	266.5	8.8	1205	US-09-959-897-1	Sequence 1, Appli
10	264.5	8.7	1155	US-09-970-516-1	Sequence 1, Appli
11	257.5	8.5	1533	US-09-205-258-90	Sequence 90, Appl
12	247.5	8.2	1149	US-09-970-516-5	Sequence 5, Appli

13	228	7.5	1611	4	US-09-248-796A-1756	Sequence 1756, Ap
14	217.5	7.2	1764	4	US-09-949-016-3940	Sequence 3940, Ap
15	210.5	7.0	1875	4	US-09-614-221A-399	Sequence 399, App
16	199.5	6.6	2462	4	US-09-620-312D-796	Sequence 796, App
17	189.5	6.3	1012	4	US-09-270-767-12677	Sequence 12677, A
18	153	5.1	942	4	US-09-107-532A-3500	Sequence 3500, Ap
19	153	5.1	1239	4	US-09-603-208A-263	Sequence 263, App
20	144	4.8	959	4	US-09-270-767-6490	Sequence 6490, Ap
21	144	4.8	959	4	US-09-270-767-21772	Sequence 21772, A
22	140	4.6	7195	4	US-09-949-016-12897	Sequence 12897, A
23	140	4.6	7198	4	US-09-949-016-15682	Sequence 15682, A
24	132.5	4.4	441529	3	US-09-103-840A-1	Sequence 1, Appli
25	129	4.3	3120	4	US-09-902-540-5505	Sequence 5505, Ap
26	129	4.3	3122	4	US-09-902-540-469	Sequence 469, App
27	126	4.2	10717	4	US-09-902-540-991	Sequence 991, App
28	125	4.1	485	4	US-09-270-767-31476	Sequence 31476, A
29	123.5	4.1	915	4	US-09-134-000C-1961	Sequence 1961, Ap
30	119.5	4.0	3089	4	US-09-634-238-188	Sequence 188, App
31	119	3.9	17612	3	US-08-911-853-29	Sequence 29, Appl
32	119	3.9	17612	3	US-09-479-409-29	Sequence 29, Appl
33	119	3.9	17612	3	US-09-479-453-29	Sequence 29, Appl
34	119	3.9	4403765	3	US-09-103-840A-2	Sequence 2, Appli
35	118.5	3.9	11575	4	US-09-938-956-5	Sequence 5, Appli
36	118	3.9	912	4	US-09-107-532A-1479	Sequence 1479, Ap
37	118	3.9	8971	4	US-09-949-016-15668	Sequence 15668, A
38	117	3.9	2963	3	US-09-232-200-60	Sequence 60, Appl
39	117	3.9	2963	3	US-09-232-197-60	Sequence 60, Appl
40	117	3.9	2963	3	US-09-232-201-60	Sequence 60, Appl
41	117	3.9	2963	4	US-09-232-195-60	Sequence 60, Appl
42	116.5	3.9	885	4	US-09-583-110-1079	Sequence 1079, Ap
43	116.5	3.9	948	4	US-09-107-433-518	Sequence 518, App
44	115	3.8	2469	4	US-09-252-991A-5908	Sequence 5908, Ap
45	115	3.8	3135	4	US-09-252-991A-5922	Sequence 5922, Ap

ALIGNMENTS

RESULT 1
US-09-774-528-148
; Sequence 148, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_Fl_genes Version 2.0
; SEQ ID NO 148
; LENGTH: 4432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129) .. (2817)
US-09-774-528-148
Alignment Scores:

Pred. No.:	6.52e-279	Length:	4432
Score:	2931.50	Matches:	548
Percent Similarity:	98.21%	Conservative:	2
Best Local Similarity:	97.86%	Mismatches:	7
Query Match:	96.91%	Indels:	3
DB:	4	Gaps:	1
US-10-631-958-11 (1-562) x US-09-774-528-148 (1-4432)			
Qy	6	GlyProAlaProLeuGly-----ValArgAlaProProAlaTrpArgThrSerPro	22
Db	1135	GGTTCCCGCCCTATAGACACAGACAGTCACAGGAGCTGGCAGGCTCTGGCGACAGCCCG	1194
Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpVallys	42
Db	1195	CGCGCGGAGATGGGGCGCGCGGGCGCGGAGCCGCTGCATCCGTGCTGTGGGTGAAG	1254
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer	62
Db	1255	CAGCAGCGCTCGCGCGTGAGCCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGCGGAGC	1314
Qy	63	ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle	82
Db	1315	CCGGGGCCCGAGCGCGCGCCCCCGCGCGGATCGCTCTGTGCTGTATCTTGAGATC	1374
Qy	83	IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys	102
Db	1375	ATCGCCGTGGAGAAACACAGACGTTACGGGAAACATCAGGCGCTGGAAATGCGCAGAA	1434
Qy	103	MetGluLysProTyrAlaPheThrValHisCysVallysArgAlaArgHisArgTrp	122
Db	1435	ATGGAAGAGCGTTACCGCTTTTACAGTTTACCTGTGTAAAGAGAGCAGCAGCGCCTGG	1494
Qy	123	LysTrpAlaGlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGln	142
Db	1495	AAGTGGCGCAGGTGACTTCTGTGTCCAGAGAGCAGCTGTGTCTACTTGTGCTGCAG	1554
Qy	143	ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle	162
Db	1555	ACCTCGCGGAGATGCTGGAGAAGCTGACGCTCCAGACCAAGCATTTTACTGGTATTATC	1614
Qy	163	AsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu	182
Db	1615	AACCCGTTTGGAGGAAAGGACAAAGCAAGCGGATATGAAAGAAAAAGTGGCACCAC	1674
Qy	183	PheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLys	202
Db	1675	TTACCTTTAGCCTCCATCACCACCTGACATCATCGTTTACTGAACATGCTTAATCAGG	1734
Qy	203	GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyYasp	222
Db	1735	GAGACTCTGTATGAGATTAAATAGACAAATACCGGCATCGTCTGTCTGGCGGAGAT	1794
Qy	223	GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal	242
Db	1795	GGTATGTTACGAGAGTGTCTCAGCGTCTGATTGGGAGGACGACAGAGAGCGCGGGTC	1854
Qy	243	AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIlePro	262
Db	1855	GACCAGAACCAACCCCGGGCTGTGTGTGTCCCGAGTAGCCTCCCGATTGGAAATCAT	1914
Qy	263	AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer	282
Db	1915	GCAGGGTCAACGGACTGGTGTGTATTCTCCACCGTGGGCACGACGACGAGAACTCG	1974
Qy	283	AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHisAsn	302
Db	1975	CGCGTGCATATCGTTGTTGGGAGCTCGGTGGCCATGGATGTGTCTCAGTCCACCAAC	2034
Qy	303	SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle	322
Db	2035	AGCACATCTCTTCCGTACTCCGTGTCCCTGTGGGTACGGCTTCTACGGGAGATCATC	2094

Qy	323	LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys	342
Db	2095	AAGCAGACTGAGAAGAAACGGTGGTGGCTTTGCCAGATACGACTTTTTCAGGTTTAAAG	2154
Qy	343	ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr	362
Db	2155	ACCTTCTCTCCACCACCTCTATGGAGGGACAGTGTCTTCTCTCCCTGCACAAACACGC	2214
Qy	363	ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer	382
Db	2215	GTGGATCTCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCAAGC	2274
Qy	383	LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp	402
Db	2275	AAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAG	2334
Qy	403	ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer	422
Db	2335	GTGGAGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC	2394
Qy	423	CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer	442
Db	2395	TGTCTTGTCCCGAGGCCCGAGGGCTCTCCCGGCTGCCACTTGGAGAGCGGTCT	2454
Qy	443	SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg	462
Db	2455	TCTGACCTCATCTCTATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAG	2514
Qy	463	HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys	482
Db	2515	CACACCACAGCAGGAGCCAGTTTGACTTCACCTTTTGTGAAGTTTATCGCGTCAAGAA	2574
Qy	483	PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLys	502
Db	2575	TTCCAGTTTACGTGGAAGCACATGGAGGATGAGCAGCAGCCCTCAAGAGGGGGGAG	2634
Qy	503	LysArgPheGlyHisLysCysSerSerHisProSerCysCysThrValSerAsnSer	522
Db	2635	AAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCTCTGCTGCTGCACCGTCTCCACAG	2694
Qy	523	SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys	542
Db	2695	TCTTGGAACTGCGACGGGAGGTCTTGCACAGCCCTGCCATCGAGGTTCAGAGTCCACTGC	2754
Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer	562
Db	2755	CAGCTGTTTCGATCTCTTGCAGGAGGAAATTGAAGAGNATCCGAAGCCAGACTCACAGC	2814

RESULT 2

US-09-270-767-14306
; Sequence 14306, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 14306
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14306

Alignment Scores:
Pred. No.: 4.55e-70 Length: 2064
Score: 813.00 Matches: 207
Percent Similarity: 46.33% Conservative: 102
Best Local Similarity: 31.03% Mismatches: 218
Query Match: 26.88% Indels: 141
DB: 4 Gaps: 17


```
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15155
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15155

Alignment Scores:
Pred. No.: 4,97e-48 Length: 1084
Score: 585.50 Matches: 125
Percent Similarity: 51.14% Conservative: 55
Best Local Similarity: 35.51% Mismatches: 103
Query Match: 19.36% Indels: 69
DB: 4 Gaps: 6

US-10-631-958-11 (1-562) x US-09-270-767-15155 (1-1084)

Qy 154 ArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArg 173
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1044 CGAGTGCCTCCCTTTTGGTCTTTATTAACCCCTATGAGGTGCGAAGCGGGAGCTCAG 985

Qy 174 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle 193
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
984 ACCTATGAGCGCATGTGAGACCTATTTCAGCTTCCGCGGTAGAGCGCCACGTGTATC 925

Qy 194 ValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyr 213
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
924 ACCACTCAGAGGGCAACCAAGTGAAGACATATCTTCTGAGCCATGATCTGGGAGTATAC 865

Qy 214 AspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIle 233
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
864 GATCGGTTTCTGTGTGCGAGGGCATGGCACCTAGTACGAGAGGTGCATCAACGGACTGATA 805

Qy 234 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 253
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
804 TTCCGTCAATGTCGAGAGTTGGACTGGAGCAACAGCGGCCACCATACATTCCAAGACCG 745

Qy 254 SerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThr 273
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
744 ---GCTTCGCAGTGGGTGTATTCCCGCTGGCAGCACCGACCATTCGCTATAGTATG 688

Qy 274 ValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla 293
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
687 CACGCACCGCGGATGTGAGCAGCGGCTATCCATGTGATCTGGGCCACGATCGGGGA 628

Qy 294 MetAspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeu 313
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
627 TTGGATGTGTGAGTGTGAGCAATGGCCAGTCCCTGCTCAGATTCTGTGCCAGTGTCTG 568

Qy 314 GlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTfTrpLeuGlyLeu 333
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
567 AGCTACGGGTACCTGGCGGATGTGGCAGCCAGACCGAGAACTACCGCTGGATGGGACCG 508

Qy 334 AlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThr 353
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
507 CGCCGGTACGAGTACAGTGGCGTCAAGCCCTTCCTGAATAATCCGGCTATGACGCCGAA 448

Qy 354 ValSerPheLeu-----ProAlaGlnHisThrValGly 364
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
447 CTGAGATGTTAGAGAGCCCGATCTTCTACTGACACCGCGCTGGAGGACATTCGCGAG 388

Qy 365 SerProArgAsp-----ArgLysProCysArgAlaGlyCys 376
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
387 AGTCCGATAGTGTGTCTGCTGGAGAGTCAGTGCCATCCGTCTGTATGCAATTCG 328

Qy 377 PheValCys-----ArgGlnSerLysGlnGlnLeuGlu 388
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
327 CAGCGCTGCAGCTTCGCCAGCAGCATACAGGAACAGCGATCCTCATTTGTTCCAAAGAG 268
```

RESULT 4

```
US-09-270-767-30448
; Sequence 30448, Application US/09270767
; Patent NO. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30448
; LENGTH: 901
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-30448

Alignment Scores:
Pred. No.: 3,45e-20 Length: 901
Score: 302.50 Matches: 85
Percent Similarity: 41.24% Conservative: 35
Best Local Similarity: 29.21% Mismatches: 82
Query Match: 10.00% Indels: 89
DB: 4 Gaps: 9
```

US-10-631-958-11 (1-562) x US-09-270-767-30448 (1-901)

```
Qy 342 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu----- 357
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2 AAGSCCTTCTGTAATATCGCGCTATGACGCCGAACTGAGAATGTTAGAAGAGCCCGAT 61

Qy 358 -----ProAlaGlnHisThrValGlySerProArgAsp----- 368
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
62 CTTCTACTGACCGCGCTGGAGACATTCGCGAGATCCGATAGTGTGTGCTCGCTG 121

Qy 369 -----ArgLysProCysArgAlaGlyCysPheValCys----- 379
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
122 GGAGAGTCAGTGCATCCGCTGCTATGCCAATTCGACGCGCTGCAGCTTCGCCAGCAGC 181

Qy 380 -----ArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrGly 396
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
182 ATACAGGAACAGCGATCCTCATTTGTTTCATCCAAGAGGAATCTAAAGAGGCA----- 232

Qy 397 LeuGluAlaAlaGluAspValGlu----- 404
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
233 ---GAGCCCAATCAGCAGGTAGAAACAGAGACTTCTCATCTAGCCCGCAGTGAAGACGA 289

Qy 404 ----- 404
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
290 CTCTGAGGCTCGTCCGCGTCCAGGCAATCTTCGATTGCCACTGGCTCCATTTCATCA 349

Qy 405 -----GluTrpGlnValValCysGlyLysPheLeuAlaIleAsn 417
```

```

Db      350 ATGAGAACCTCGGCAACGATCAGTGAAGGTTGCGGGCAATTTCTTATGATCGC 409
Qy      418 AlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHis 437
Db      410 GCGCGAAACATAACCTGCGCTGCGCAGGAGTCCCAATGGCATCTCCGTTTACAGTCAT 469
Qy      438 LeuGlyAspGlySerSerAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 457
Db      470 CTGGGTGATGGTTGCTGACCTGATTTCTCGTGAAGAACCTCACTGCTCAACAAACGTG 529
Qy      458 ArgPheLeuLeuLeuArgHisThrAsnGlnGln---AspGlnPheAspPheThrPheValGlu 476
Db      530 CGTTTCTGCTCAACACGCGCGGAGAGTGTGATATTCGCAATTTGCTTTTGTAGAG 589
Qy      477 ValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAsp 496
Db      590 GTATATCGACAAAGGAATTCGGTTTCAGAACTTTTTCGCGAGGAGGAGGACTACAGC 649
Qy      497 LeuLysGluGlyGlyLysLysArgPheGlyHisLysCys----- 509
Db      650 TTGCGAGGCTCT-----TGTACGCCGATTACACCGCCTGAA 685
Qy      510 -----SerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCysAsp 527
Db      686 GAATGACCGCCCATCTCTCC-----TCGACGGAGTTTTCAGCTGGAACGTGAT 736
Qy      528 GlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArgLeu 547
Db      737 CGAGAGGTGGTGACCGACTGGACATACCATGCGATCGCATTTGTCAGCTCATCGAGTGC 796
Qy      548 PheAlaArgGlyLeuGluGluAsnProLysPro 558
Db      797 TTCATGCGGGTCCCATTTCTTACAGCAAGCCC 829

```

RESULT 5

```

US-09-970-516-3
; Sequence 3, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE OF INVENTION: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1857)
; OTHER INFORMATION:
US-09-970-516-3

```

```

Alignment Scores:
Pred. No.:      2,648-18      Length:      1857
Score:          288.50      Matches:      97
Percent Similarity: 42.74%      Conservative: 59
Best Local Similarity: 26.58%      Mismatches: 173
Query Match:      9.54%      Indels:      37
DB:              4          Gaps:      7

```

US-10-631-958-11 (1-562) x US-09-970-516-3 (1-1857)

```

Qy      18 TrpArgThrSerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSer 37
Db      23 TGGCTGCCAGCACCCCGCTCTCCATGGCGAGTTGGCTCTTACCCAGCCCGAGGCCAC 82
Qy      38 ValLeuTrpValLysGlnGlnArg-----CysAlaValSerLeu 50

```

RESULT 6

US-09-817-676A-13

```

Db      83 GCTTTGCCCTCACCTTACATCGAGGCCCTGACATACAGCGGCTGCGCCCCCAACCTG 142
Qy      51 GluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyProGlyAlaGlyAlaPro 70
Db      143 AAGCAGGCCCGGG---GTGGCCTGGTCCCGTGGCCGAGGTCTCAGGCTGCTGCACCC 199
Qy      71 GlyAlaAlaCysSerValProValSerGluLeuLeuAlaValGluGluThrAspVal 90
Db      200 ---TGCAGAACGCGACCCCTCAGACTCAGCGGCTACTTCTGCATCTAC-ACCTACCCCT 255
Qy      91 HisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysProTyrAlaPheThr 110
Db      256 CGGGGCGCGCGCGGGCGCGAGCCACTCGACCTTCCGGGCGAGATGGGGCGGCC 315
Qy      111 ValHisCysValLysArgAlaArgHisArgTrpLysTrpAlaGlnValThrPheTrp 130
Db      316 ACCTACGAAGAGAACCGTGGCGAGCCAGCGCTGGGCCACTGCGCCTCACCTGCTG--- 372
Qy      131 CysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLys 150
Db      373 -----CTCCGAGGACTGCTCACTCCGCGGATGGGAGATCACCCCTGAC 417
Qy      151 LeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlyGln 170
Db      418 CTGCTACTCGCGCGCGCGGTTGCTTCTATTGTCATCCCTTTGGGGGTCGGGGCGCTG 477
Qy      171 GlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThr 190
Db      478 GCCTGGCAGTGGTGTAGAACCCAGCTGCTTCCCATGATCTCTGAAGCTGGGCTGCTCTTC 537
Qy      191 AspileleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIle 210
Db      538 AACCTCATCCAGACAGAACAGACAGAACACGCCCGGAGCTGGTCCAGGGGCTGAGCCTG 597
Qy      211 AsplysTyrAspGlyIleValCysValGlyCysGlyMetPheSerGluValLeuHis 230
Db      598 AGTGAGTGGGATGGCATCGTCACGGTCTCGGGAGACGGGCTGCTCCATAGAGTCTGTAAC 657
Qy      231 GlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaVal 250
Db      658 GGGCTCTAGATCGC-----CCTGACTGGGAGGAGCTGTG 693
Qy      251 LeuValProSerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCys 270
Db      694 AAGATGCGCT-----GTGGGCATCTCCCTCGCGCTCGGCAACGCGCTGGCC 741
Qy      271 TyrSer-----ThrValGlyThrSerAspAlaGluThr 281
Db      742 GGAGCAGTGAACACGACGCGGGGATTTGAGCCGCTTGGGCGCTCGACCTGTTGCTCAAC 801
Qy      282 SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHis 301
Db      802 TGCTCACTGTGTGTCGCGGGGTGGTGGCCACCACCTGACCTGCTCTCGTGACCGCTG 861
Qy      302 AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspile 321
Db      862 GCCTCGGGCTCCCGCTGTTTCTCTCTGCTGCTGGCGCTTGGTGTGTCAGATGTG 921
Qy      322 IleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 341
Db      922 GATATCCAGAGCAGCGCTTTCAGGGCGCTTGGGCGAGTGCCTGCTTTCACACTGGGACCGGTG 981
Qy      342 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 361
Db      982 CTGGGCTCGCCACACCTGACACCTACCGGAGCGCTCTCTTACCTTCCCGCCACCTGTG 1041
Qy      362 ThrValGlySerPro 366
Db      1042 GAACCTGCTCGCCC 1056

```

Sequence	Accession	Species	Gene	Protein	Length	Score	Percent Similarity	Best Local Similarity	Query Match	DB
Sequence 13, Application US/09817676A	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
Patent No. 6800470	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
GENERAL INFORMATION:	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
APPLICANT: Spiegel, Sarah	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
APPLICANT: Kohama, Takafumi	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning, Expression and Methods of Use Thereof	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
TITLE OF INVENTION: Expression and Methods of Use Thereof	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
FILE REFERENCE: 00170/HG	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
CURRENT APPLICATION NUMBER: US/09/817,676A	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
CURRENT FILING DATE: 2001-03-26	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
PRIOR APPLICATION NUMBER: US 60/194,318	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
PRIOR FILING DATE: 2000-04-03	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
NUMBER OF SEQ ID NOS: 15	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
SOFTWARE: Patent In Ver. 2.0	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
SEQ ID NO 13	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
LENGTH: 2380	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
TYPE: DNA	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
ORGANISM: Homo sapiens	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
FEATURE:	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
NAME/KEY: CDS	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
LOCATION: (7)..(1860)	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
PUBLICATION INFORMATION:	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
TITLE: Molecular cloning and functional characterization of a	US/09817676A	Human	Protein	Protein	2380	3.91e-18	42.74%	26.58%	9.54%	US-10-631-958-11 (1-562) x US-09-817-676A-13 (1-2380)
TITLE: novel mammalian sphingosine kinase type 2 isoform	US/09817676A	Human								

Score: 280.50 Matches: 143
Percent Similarity: 37.82% Conservative: 82
Best Local Similarity: 24.03% Mismatches: 223
Query Match: 9.27% Indels: 148
DB: 4 Gaps: 27

US-10-631-958-11 (1-562) x US-09-949-016-1155 (1-1783)

QY 2 GluAlaAlaAsnGlyProAlaProLeu-----GlyValArgAlaProProAlaTrpArg 19
DB 2 GAGGAGCGAGCGCGGAGTCCGCTCAGCGGGCGCTCCAGTCCCTCAGACGTGGGCT 61
QY 20 -----ThrSerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGlu 33
DB 62 GAGCTTGGGACGAGCTGGTCCGCCCGCCAGGCCACTGTAGGGAACGGGGTGGCGCTCC 121
QY 34 ProLeuGlnSerValLeuTrpValys-----GlnGlnArgCysAlaValSerLeuGlu 51
DB 122 CCAGCAAAACCGGACCGACTGGGTCCAGCGCGCGGAGGAATGACACCGGTGCTCTACAG 181
QY 52 ProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyProGlyAlaGlyAlaProGly 71
DB 182 CCA-----CGGCT-CCGGGC---GGGGAAGGCGAGCCCCAC 213
QY 72 AlaAspAlaCysSerValProValSerGluIleAlaValGluLeuThrAspValHis 91
DB 214 AGCGCGCCCTGGACCGCGCTGGGC-----AGCACCGAT----- 249
QY 92 GlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysProTyrAlaPheThrVal 111
DB 250 ---AGGAGCTCAAGCGAGGACCGCCCGCCAGCGGCGGCGCC----- 291
QY 112 HisCysValLysArgAlaArgHisArgTrpLysTrpAlaGlnValThrPheTrpCys 131
DB 292 -----ACAGCGCCAGGACCGCCCTGGCAG 315
QY 132 ProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLys--- 150
DB 316 CGGGAGCGCGGGTC-----GAGGTTATGATCCACGC 348
QY 151 -----LeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 165
DB 349 GCGGCGCCCGGGGGTGTCCCGCGGCTCGCCGCTGCTGGTGTCTGTAACCGCGC 408
QY 166 GlyLysGlyGlnGlyLeuArgLysValAlaProLeuPheThrLeu 185
DB 409 GCGGCAAGGCAAGCGCTTCAGCTCTCCGAGTCACTGCGAGCGCCCTTTGGCTCAG 468
QY 186 AlaSerIleThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 205
DB 469 GCTGAATCTCTTCCACGCTGATGCTCACTGAGCGCGGAACCCAGCGCGGAGCTGGT 528
QY 206 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyAspGlyMetPhe 225
DB 529 CGGTCCGAGGAGCTGGCGGCTGGACGCTCTGGTGTCTATGTCTGGAGCGGCTGATG 588
QY 226 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 245
DB 589 CACGAGGTGGTGAACGGGCTCATGAGCGGCTGACTGGGACCGCCATCCAGAAG-- 645
QY 246 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAlaGlySer 265
DB 646 -----CCCTGTGTAGCTCTCCACGAGGCTCT 672
QY 266 ThrAspCysValCys-----TyrSerThrValGlyThrSer 277
DB 673 GCAACCGGCTGGAGCTTCTTGAAACCATATGCTGGCTATGACGAGTCAACATGAA 732
QY 278 AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla---MetAspVal 296
DB 733 GACCTCTGACCACTGACGCGTATGTGTGCGCGCGCTGCTGTACCATGAACCTG 792
QY 297 SerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGly 316

Db 793 CTGTCTTGACACCGGCTTGGGGCTGGCGCTCTTCTGTGCTCAGCTGGCTGGGCG 852
QY 317 PheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyr 336
DB 853 TTCAATGCTGATGTGGACCTAGAGAGTATCGGCGTCTGGGGAGATGCGCTTC 912
QY 337 AspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPhe 356
DB 913 ACTCTGGGCACTTCTCGCGCTCTGGCAGCCCTCGCACCTACCGCGCGGCTGCTAC 972
QY 357 LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 376
DB 973 CTCCTCTAGGAAGA---GTGGGTTCACAGACACTGCTCTCCCT----- 1014
QY 377 PheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGly 396
DB 1015 GTTGTGTCCAGCAGGCGCGGTAGATGCACAC-----CTTGTGCCA 1056
QY 397 LeuGluAlaAlaGluAspVal---GluGluTrpGlnValValCysGlyLys----- 412
DB 1057 CTGGAG-----GAGCAGTGCCTCTCTACTGGACAATGGTGCACGAGGACTTGTG 1110
QY 413 ---PheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly 431
DB 1111 CTAATCTGGCAGCTGCTGCACCTGCACCTGGGAGTGAATGTTGTGTGCACCATGGGC 1170
QY 432 LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuLys--- 450
DB 1171 CGCTGTGCAGCT-----GGCGTCATGCATCTGTCTACGTGCGGCGGGA 1215
QY 451 CysSerArgPheAsnPheLeuArgPheLeuIle-----ArgHisThrAsn 465
DB 1216 GTGTCTGTGCATGCTGCTGCGCTTCTCTCTGCTCCATGGAGAGGCGGAGCATATG--- 1272
QY 466 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 485
DB 1273 -----GAGTATGATGCTCTACTGTATATGTGCCCGTGTGCTTCCGCTCCGC--- 1320
QY 486 ThrSerLysHisMetGluAspGlySerAspLeuLysGluGlyLysLysArgPhe 505
DB 1321 -----TTGGAGCCCAAGAT-----GGGAAAGGTGTGTTT 1350
QY 506 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 525
DB 1351 GCA----- 1353
QY 526 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 545
DB 1354 GTGATGGGAATGTAGTGTAGCGAGCGCTGCGGGCCAGGTGCACCCAACTACTTTC 1413
QY 546 ArgLeuPheAlaArgGlyIleGluGluAsnPro-----LysPro 558
DB 1414 TGGATGTGCTAGTGTGCTGGAGCCCGCCCGCCAGCTGGAGGCC 1458

RESULT 8

US-09-817-676A-11
; Sequence 11, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; TITLE OF INVENTION: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; PRIOR APPLICATION NUMBER: 2001-03-26
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2698

```

; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (387)..(2237)
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245448
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-11

Alignment Scores:
Pred. No.: 4,04e-16 Length: 2698
Score: 269.00 Matches: 98
Percent Similarity: 41.64% Conservative: 59
Best Local Similarity: 25.99% Mismatches: 159
Query Match: 8.89% Indels: 62
DB: 4 Gaps: 8

US-10-631-958-11 (1-562) x US-09-817-676A-11 (1-2698)

Qy 44 GlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrrpArgSerPro 63
Db 427 CAATCTGCAGCGGAGTTTGTTCTACCCGCCCAAGCCAGCGGTTGCCCTCACCC 486

Qy 64 Gly-----ProGlyAlaGlyAlaProGlyAlaAsp 73
Db 487 TCACAACAAGCCCTACACATACAGCGACTACGCCCAAGCCAGAAAGCCCGCCGAG 546

Qy 74 AlaCysSerValProValSerGluLeuLele-----AlaVal 85
Db 547 ATGGTCTA-GTCTCTCTGGATGAGGTCTCGGGCTGTGGCCACCTCGACAGCGCGTAGCCCC 605

Qy 86 GluGluThrAspValHis-----GlyLysHisGlnGlySer 97
Db 606 GAGGACACTGCAGCCCTTCTTCATCTACACTACCACGTGGCGCGTCGAGGGGGCGG 665

Qy 98 GlyLeuTrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAla 117
Db 666 CGCAGAGCTACCGGACCTTCGCGGGCGGATGGGGCCACCACCTATGAGGAGAATCGTGA 725

Qy 118 ArgArgHisArgTrrpLysTrpAlaGlnValThrPheTrrpCysProGluGlnLeuCys 137
Db 726 GAGGCCCGAGCGCTGGGCCACTGCCTCACGTGCTC-----CTCCGA 767

Qy 138 HisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHis 157
Db 768 GGAATCGCTCTGFCAGGGGACCCAGGAATAACCCCTGAATTGCTCCCGCCAGGCCAGG 827

Qy 158 LeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTrrpGluArg 177
Db 828 CTGCTCATATTGGTCAATCCCTTTGGGGGGCGGGCGCTGGCTGCAGCGGTGTATGGAC 887

Qy 178 LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleValThrGluHis 197
Db 888 CACGTGTGCAATGATCTCTGAAGCTGGGCTGTCTCTCAACCTCATACAGACAGAACGA 947

Qy 198 AlaAsnGlnAlaLysGluThrLeuTrrpGluLeuIleAsnIleAspLysTrrpAspGlyLeuVal 217
Db 948 CAGAACCATCGCCCGTGCAGCTGTCGAGGGGTTAAAGCCCTGAGTGAGTGGGAAGGCATTGTC 1007

Qy 218 CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln 237
Db 1008 ACTGTGTCTGGAGACGGGCTGCTTTACGAGGTGCTGTAATGGGCTCTCTGTATCGG----- 1061

Qy 238 ArgSerAlaGlyValAspGlnAenHisProArgAlaValLeuValProSerSerLeuArg 257

```

```
QY 154 ArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyLeuGlyGlnGlyLysArg 173
Db 69 CGGCCCTGCCGCTGCTGTCTGAACCCGCGCGGCAAGGCAAGCCCTTGCAG 128
QY 174 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleile 193
Db 129 CTCCTCCGGAGTCACGTGCAGCCCTTTTGGCTGAGGCTGAATCTCCTTCACCGCTGATG 188
QY 194 ValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyr 213
Db 189 CTCACGTAGCGCGGCAACACCGCGGAGTGTGTGGCTCGGAGGAGCTGGCGCGCTGG 248
QY 214 AspGlyIleValCysValGlyAspGlyMetPheSerGluValLeuHisGlyLeuIle 233
Db 249 GACGCTCTGGTGTCTGTCGACCGGCTGATGCAGAGGTGGTGAACGGGCTCATG 308
QY 234 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 253
Db 309 GAGCGGCTGACTGGGAGACCCCATCCAGAG----- 341
QY 254 SerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCys----- 270
Db 342 -----CCCTGTGTAGCTCCACAGAGCTCTGGCAACGCGCTGGCAGCTTCCCTTG 392
QY 271 -----TyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 285
Db 393 AACCATTTATGCTGTGATGAGCAGCTCACCAATGAAGACCTCTGACCACTGACGGCTA 452
QY 286 IleValValGlyAspSerLeuAla---MetAspValSerSerValHisAsnSerThr 304
Db 453 TTGCTGTCCGCGCGCTGCTGTACCCATGAACCTGCTCTCTGCACACGCGTTCGGGG 512
QY 305 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAsp 324
Db 513 CTGCGCCTCTTCTGTGCTCAGCCTGCGCTGGCTGCTCATTTGCTGATGTGGACCTAGAG 572
QY 325 SerGluLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 344
Db 573 AGTGAGAGTATCGCGCTGTGGGGAGAGTGGCTTCACTCTGGGCACCTTCTCGCTGCTG 632
QY 345 LeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGly 364
Db 633 GCAGCTTGGCAGCTTACCGCGCGAGTGGCTTACCTCCCTGTAGGAGA---GTGGGT 689
QY 365 SerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGln 384
Db 690 TCCAAGACACCTGCCTCCGCC-----GTTGTGCTCCAGCAGGCGCCGCTA 734
QY 385 GlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal--- 403
Db 735 GATGCACAC-----CTTGTGCCACTGGAG-----GAGCCAGTGCCTCC 770
QY 404 GluGluTyrGlnValValCysGlyLys-----PheLeuAlaIleAsnAlaThr 419
Db 771 TCTACTGGACAGTGGTCCCGACGAGACTTTGTGTAGTCTGCTGGCACTGCTGCACCTG 830
QY 420 AsnMetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 439
Db 831 CACCTGGGCGAGTGTCTGCTGCACCCATCCGCGCGCTGTGTCAGCT----- 878
QY 440 AspGlySerSerAspLeuIleLeuArgLys---CysSerArgPheAsnPheLeuArg 458
Db 879 ---GGCGTCATGCATCTGTTCACGTGGCGGGGAGTGTCTCGTGCATGCTGCTGCTGCGC 935
QY 459 ---PheLeuIleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluVal 477
Db 936 CTCCTTCTGGCCATGAGAGGCGAGGATATGAGATGATGATGATGATGATGATGAT 995
QY 478 TyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu 497
Db 996 GTGCCCGTGGTGCCTTCCGC-----TTGGAGCCCAAGGAT----- 1031
QY 498 LysGluGlyGlyLysArgPheGlyHisIleCysSerSerHisProSerCysCysCys 517
```

```
Db 1032 -----GGGAAGGTATGTTTGA----- 1049
QY 518 ThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGlu 537
Db 1050 -----GTGGATGGGAAATTGATGTTAGCGAGCCGCTGCAG 1085
QY 538 ValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnPro--- 556
Db 1086 GGCAGGTGACCCCAAACTACTTCTGATGTCAGCGTTCGCTGGAGCCCCCGCCAGC 1145
QY 557 ---LysPro 558
Db 1146 TGAAGGCC 1154
RESULT 10
US-09-970-516-1
; Sequence 1, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE OF INVENTION: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1155)
; OTHER INFORMATION:
US-09-970-516-1
Alignment Scores:
Pred. No.: 2,92e-16 Length: 1155
Score: 264.50 Matches: 107
Percent Similarity: 41.37% Conservative: 68
Best Local Similarity: 25.30% Mismatches: 169
Query Match: 8.74% Indels: 79
DB: Gaps: 16
US-10-631-958-11 (1-562) x US-09-970-516-1 (1-1155)
QY 154 ArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArg 173
Db 37 CGGCCCTGCCGCTGCTGTCTGAACCCGCGCGGCAAGGCAAGCCCTTGCAG 96
QY 174 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleile 193
Db 97 CTCCTCCGGAGTCACGTGCAGCCCTTTTGGCTGAGGCTGAATCTCCTTCACCGCTGATG 156
QY 194 ValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyr 213
Db 157 CTCACGTAGCGCGGGAACACCGCGCGAGCTGTGTGGCTCGGAGGAGCTGGCGCGCTGG 216
QY 214 AspGlyIleValCysValGlyAspGlyMetPheSerGluValLeuHisGlyLeuIle 233
Db 217 GACGCTCTGGTGTGTCATCTCTGGAGACGGGCTGATGCACGAGGTGGTGAACGGGCTCATG 276
QY 234 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 253
Db 277 GAGCGGCTGACTGGGAGACCGCATCCAGAG----- 309
QY 254 SerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCys----- 270
Db 310 -----CCCTGTGTAGCTCCACAGAGCTCTGCAACGCGCTGGCAGCTTCCCTTG 360
QY 271 -----TyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 285
```



```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 90
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1522)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1527)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-90

Alignment Scores:
Pred. No.: 2,258-15 Length: 1533
Score: 257.50 Matches: 113
Percent Similarity: 41.75% Conservative: 64
Best Local Similarity: 26.65% Mismatches: 167
Query Match: 8.51% Indels: 80
DB: 4 Gaps: 18

US-10-631-958-11 (1-562) x US-09-205-258-90 (1-1533)

Qy 154 ArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArg 173
Db 109 CGGCCCTCCGGTGNCTGGTCTGTAACCCGCGCGGCAAGGCGGCAAGCCCTTGCAG 168
Qy 174 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIle 193
Db 169 CTCITCCGGAGTCACGTGCAGCCCTTTTGGCTGAGGCTGAATCTCTTCACGCTGATG 228
Qy 194 ValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyr 213
Db 229 CTCACGTAGCGCGGAACCAACGCGCGGAGTCTGGTGGTGGAGGAGCTGGCGCGCTGG 288
Qy 214 AspGlyIleValCysValGlyAspGlyMetPheSerGluValLeuHisGlyLeuIle 233
Db 289 RACGCTCTGGTGTCTGTTGGACCGGCTGATGCACGAGGTGGTGAACGGGCTTCAT 348
Qy 234 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 253
Db 349 GGA-----GCGGCTGACTGGGAGACCCCATCCAGAGCC-----CCT 387
Qy 254 -SerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCys----- 270
Db 388 GTGTAGCCTC-----CCAGCAGGCTCTGGCAACGCGCTGGCAGCTTCCTT 432
Qy 271 -----TyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHi 285
Db 433 RAACATTATGCTGGCTATGACAGGTCACCAATGAAGACCTCTCGACCACTGCACGCT 492
Qy 285 sIleValValGlyAspSerLeuAla---MetAspValSerValHisHisAsnSerTh 304
Db 493 ATTGCTGTGGCGCGGCTGTGTACCATGAACTGTCTGTCTGTGCACAGGCTTCGGG 552
```

```
Qy 304 rLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAs 324
Db 553 GCTGCGCCTCTTCTCTGTCTGCTCAGCCTGGCTGGGCTTCATTGCTGATGGACCTAGA 612
Qy 324 pSerGluLysLysArgTrpLeuGlyLeuAlaAArgTyrAspPheSerGlyLeuLysThrPh 344
Db 613 GAGTGAGAGTATCGGCTCTGGGGAGATGCGGCTTCACTCTGGGCACCTTCCTGCGCTCT 672
Qy 344 eLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGl 364
Db 673 GGCAGCCTCGCACCTACCGCGCGCAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729
Qy 364 ySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGl 384
Db 730 TTCAAGACACCTGCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 774
Qy 384 nGlnLeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal-- 403
Db 775 AGATGCACAC-----CTTGTGCCACTGGAG-----GAGCCAGTGCC 810
Qy 404 -GluGluTrpGlnValValCysGlyLys-----PheLeuAlaIleAsnAlaTh 419
Db 811 CTCTCACTGGACAGTGGTGGCCGACGAGGACTTGTGTAGTCTGCTGCTGCTGCTGCTGCT 870
Qy 419 rAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGl 439
Db 871 GCACCTGGGAGTGAGATGTTTGTGCACCCATGGCGCGCTGTGCAGCT----- 919
Qy 439 yAspGlySerSerAspLeuIleLeuIleArgLys---CysSerArgPheAsnPheLeuAr 458
Db 920 ----GGCGTATGATCTGTCTACGTGGGGGGGAGTGTCTGCTGCCATGCTGCTGCTGCG 975
Qy 458 g---PheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluVa 477
Db 976 CCTCTTCTGGCCATGGAGAAGGCGACGATATGAGTATGAAATGCCCTACTTGGTATA 1035
Qy 477 lTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLe 497
Db 1036 TGTGCGCGTGTGCTGCGCTTCCGC-----TTGAGCCCAAGGAT----- 1072
Qy 497 uLysGluGlyGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCy 517
Db 1073 -----GGGAAAGGTGTGTTGCA----- 1090
Qy 517 sThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGl 537
Db 1091 -----GTGGATGGGAATTGATGTTAGCGAGGCGGTGCA 1125
Qy 537 uValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnPro-- 556
Db 1126 GGGCAGGTGCACCAAACTACTTCTGGATGGTTCAGCGGTTCGCTGAGCCCCCGCCAG 1185
Qy 557 ----LysPro 558
Db 1186 CTGGAAGCCC 1195

RESULT 12
US-09-970-516-5
; Sequence 5, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Mus musculus
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1149)
; OTHER INFORMATION:
US-09-970-516-5

Alignment Scores:
Pred. No.: 1,38e-14 Length: 1149
Score: 247.50 Matches: 100
Percent Similarity: 41.65% Conservative: 67
Best Local Similarity: 24.94% Mismatches: 179
Query Match: 8.18% Indels: 56
DB: 4 Gaps: 13

US-10-631-958-11 (1-562) x US-09-970-516-5 (1-1149)

Qy 151 LeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyGln 170
Db 28 CTGCTCCACCGCCATGCAGAGTGTGTGCTGAACCCCGGGTGGCAAGGCAAG 87
Qy 171 GlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThr 190
Db 88 GCTCTGAGCTCTCCAGAGCGGTGTGCAGCCCTCTCGAGGAGGAGAGATTAACCTTT 147
Qy 191 AspileIleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIle 210
Db 148 AAACGTGATACACCGAAGCAAGAACCATCCAGGAGCTGGTGTGCAGAGGAGTGTG 207
Qy 211 AspLysTyrAspGlyIleValCysValGlyCysValGlyMetPheSerGluValLeuHis 230
Db 208 GGTCACCTGGGACGCCCTGGCAGTCATGTCGGTGTGTCGTGATGATGAGGTGGTGAAT 267
Qy 231 GlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaVal 250
Db 268 GGGCTAATGGAAACGGCCGACCTGGAGACTCCCATCCAGAA----- 309
Qy 251 LeuValProSerSerLeuArgIleGlyIleLeuProAlaGlySerThrAspCysValCys 270
Db 310 -----CCCTGTGTAGCTCTCCCTGGAGGCTCCGCAATGCCCTGCCA 351
Qy 271 -----TyrSerThrValGlyThrSerAspAlaGluThrSer 282
Db 352 GCTTCTGTGAACCACTATGCTGGGTAGCAGAGTGAATAATGAAGACCTGCTCATCAAC 411
Qy 283 AlaLeuHisIleValValGlyAspSerLeuAla---MetAspValSerSerValHisHis 301
Db 412 TGCACACTGCTGTGTGCGCGCGGCTGTGCACCAATGAACCTGCTCCTGCACACT 471
Qy 302 AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspile 321
Db 472 GCTTCTGGGCTGCGGCTCTATTCTGTGCTCAGTCTGCTCGGGGCTTTGTTGCTGACGTG 531
Qy 322 IleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 341
Db 532 GACCTCGAGAGTGAGAAGTACAGCGCTTGGGGGAGATTCTTTTCACAGTGGGCACCTTC 591
Qy 342 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 361
Db 592 TTTGCGCTAGAACGCTCGCATCTACACGACCACTGGCTACCTTCCT----- 642
Qy 362 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 381
Db 643 ---GTAGAACTGTGGCTCTAAGAGACCCGCC----- 672
Qy 382 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu 401
Db 673 -----TCTACACTGGTGCAAGAGGGCCCGCTCGACACACACTTGTCTCTGGAGGAG 726
Qy 402 AspVal---GluGluTrpGlnValValCysGlyLys-----PheLeuAlaIle 416
Db 727 CCAGTGGCTTCTCATTTGACTGTGTGTACCAAGACAGGACTTTGTCTGGTGTGCTGTG 786
Qy 417 AsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAla 436
```

```
Db 787 CTACACACCCACCTGAGCTCCGAGCTGTTTTGCAGACCCCATGGCGCGCTGTGAGGCT--- 843
Qy 437 HisLeuGlyAspGlySerSerAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 455
Db 844 -----GGTGTATGTCATCTGTTCTAGCTAGCTGCGGGGGTGTCAAGGGCTGCG 891
Qy 456 PheLeuArg---PheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPhe 474
Db 892 CTGCTGGCGCTCTCTCTGGCCATGCAGAGGCAAGCATATGGAACCTTGACTGTCCATAC 951
Qy 475 ValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAsp 494
Db 952 CTGTTTCATGTGCCCGCTGTTTCTCCGCTCGAGCCAGAGCCAGAGGGCGTGTGTTT 1011
Qy 495 SerAspLeuLysGluGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLys 514
Db 1012 TCT-----GTGATGGAGGCTGTATGG--TATGTGAAGCTGTGAGGGCCAG----- 1057
Qy 515 CysCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerPro 534
Db 1058 -----TGCACCCAAACTACCTTTGGATGGTCTGTGTCAGCAGAGATGCCCAT---CCG 1108
Qy 535 Ala 535
Db 1109 GCC 1111

RESULT 13
US-09-248-796A-1756
; Sequence 1756, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1756
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1756

Alignment Scores:
Pred. No.: 2e-12 Length: 1611
Score: 228.00 Matches: 81
Percent Similarity: 42.86% Conservative: 57
Best Local Similarity: 25.16% Mismatches: 143
Query Match: 7.54% Indels: 42
DB: 4 Gaps: 9

US-10-631-958-11 (1-562) x US-09-248-796A-1756 (1-1611)

Qy 144 LeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsn 163
Db 454 TTAACAGAAATCATATGGAAAAGTATAAATTCGCGCCCTCT---ATTTTGGTATTGATAAAC 510
Qy 164 ProPheGlyGlyLysGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPhe 183
Db 511 CCGCATGGCGCCCAAGGACAGCCAAACTATTTTACAAATAAATACTTTACCAATATTA 570
Qy 184 ThrLeuAlaSerIleThrThrAspileIleValThrGluHisAlaAsnGlnAlaLysGlu 203
Db 571 CAAGCGCTCGTGAATGTTACGTATTTTTGAACATAAATATCATGACACGCCACTGAG 630
Qy 204 ThrLeuTyrGluLeuAsnIleAspLysTyrAspGlyIleValCysValGlyAspGly 223
```

```
Db 631 ATTGCGGTGAGCTAGATGCTCAATGATTATGATATATAATGTTTGTGTTCTGGCGATCGG 690
Qy 224 MetPheSerGluValLeuHisGlyLeuLeuGlyArgThrGlnArgSerAlaGlyValAsp 243
Db 691 ATACCTCATGAGTTATCAATGGCTTCTATCTCTCCAGAT----- 732
Qy 244 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyLeuLeuProAla 263
Db 733 -----AAAGGTTTACTGCATTCACAAAATTGCGATTACCAATTACCTTGT 780
Qy 264 GlySerThrAspCysValSerThrValGlyThrSerAspAlaGluThrSerAla 283
Db 781 GGGTCAGGAACCGCTTGAGTTGATACACATGCTAGTAAATGCTTCAGTTCACACT 840
Qy 284 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHis----- 301
Db 841 CTTTATATGTTGAAAGCTCATAGAACAATAATGGATTTGATGGCTATTACCAAGGTACA 900
Qy 302 ---AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp 320
Db 901 GGAAGTGAAGATGAACGAAATGTCATTTTAAAGTCAGTGTCTCGGTATTATTGCTGAT 960
Qy 321 IleIleLysAspSerGlyLysLysArgTyrLeuGlyLeuAlaArgTyrAsp----- 337
Db 961 TCTGATATTGGACAGACATTTACGTTGGTGGTCTTATACGGTTTGAACCTGGAGTG 1020
Qy 338 -----PheSerGlyLeuLys-----ThrPheLeuSerHisHisCys 349
Db 1021 ATACAAAAGTGTCTCGGGGCAAAATATCTTGTGACCTATTGTTGAAATACAAG--- 1077
Qy 350 TyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArg 369
Db 1078 TACGATAATAATTCAGAGATTGTAATCATGTAAATGATTACTTGAGTAATAATGAT--- 1134
Qy 370 LysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnLysGluGluGlu 389
Db 1135 -----ACTGAAACGAGTTATACCATTGTCTACTGAA 1164
Qy 390 GlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal---GluGluTyrGln-Va 408
Db 1165 GAGAACTTGCAAAATACGAGCCTGATTTAGATCAACCTGTCCCAATGATTGGAGACAT 1224
Qy 408 lValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSe 428
Db 1225 ATTCTCTCAAGAAATTTCTCATAATTTGAACATCTTGTATGTTGGCAAAATGCCATTCGTG 1284
Qy 428 rProArg-----GlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuI 446
Db 1285 TCTGAGATACTCAGTTTTTC-CCTGCTGCTCTACCGAACGATGTTCAATGGACATGAT 1343
Qy 446 eLeu 447
Db 1344 TGTC 1347

RESULT 14
US-09-949-016-3940
; Sequence 3940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 3940
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3940

Alignment Scores:
Pred. No.: 2,53e-11 Length: 1764
Score: 217.50 Matches: 139
Percent Similarity: 36.99% Conservative: 80
Best Local Similarity: 23.48% Mismatches: 231
Query Match: 7.15% Indels: 144
Db: 4 Gaps: 28

US-10-631-958-11 (1-562) x US-09-949-016-3940 (1-1764)
Qy 2 GluAlaAlaAsnGlyProAlaProLeu-----GlyValArgAlaProProAlaTyrArg 19
Db 2 GAGGAGCGAGCGCGGGAGTCCGCTCCAGCGGGCGCTCCAGTCCCTCAGACGTGGGCT 61
Qy 20 -----ThrSerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGlu 33
Db 62 GAGCTTGGGACGAGCTGCGTTCGCCGCCAGCCACTGTAGGGAACGGCGTGGCGCTCC 121
Qy 34 ProLeuGlnSerValLeuTyrValLys-----GlnGlnArgCysAlaValSerLeuGlu 51
Db 122 CCAGCAACCGGACCGAGTGGTCCAGCCGCCGAGGGAATGACGCCGCTCTCTACAG 181
Qy 52 ProAlaArgAlaLeuLeuArgTyrTyrArgSerProGlyProGlyAlaGlyAlaProGly 71
Db 182 CCA-----CGGT-CGGGC---GGGGAAGGCGAGCCCGCCAC 213
Qy 72 AlaAspAlaCysSerValProValSerGluIleAlaValGluGluThrAspValHis 91
Db 214 AGCGCGCCTCGACGCGCGCTCGGC-----AGCACCGAT----- 249
Qy 92 GlyLysHisGlnGlySerGlyLysTyrGlnLysMetGluLysProTyrAlaPheThrVal 111
Db 250 ---AAGGAGCTGAAGCAGGAGGAGCGCCGCCAGCGGACGCGCC----- 291
Qy 112 HisCysValLysArgAlaArgArgHisArgTyrLysTyrAlaGlnValThrPheTyrCys 131
Db 292 -----ACAGGCCAGGAGACCCCTGGGAG 315
Qy 132 ProGluGlnLeuCysHisLeuTyrPheGlnThrLeuArgGluMetLeuGluLysLeu 151
Db 316 CGGAGCGCGCGGTC-----GAGGTTATGATCCA--- 345
Qy 152 ThrSerArgProLysHisLeuLeuVal-----PheIleAsnPro 164
Db 346 GCAGGCGCGCCCGCGGCTAGCTCCGCGCCCTTGCCTATGAGTGAAGTCTGAACCCG 405
Qy 165 PheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThr 184
Db 406 ---CGAAGGCAAGGCGAAGGCTT-CAGCTCTTCGGAGTACGTGCAGCCCTTTGGCT 461
Qy 185 LeuAlaSerIleThrThrAspIleValThrGlnHisAlaAsnGlnAlaLysGluThr 204
Db 462 GAGGCTGAATCTCTTCCAGCTGATGCTACTGAGCGCGGAAACCAACGCGGAGACTGG 521
Qy 205 LeuTyrGluIleAsnIleAspLysTyrAspGlyLeuValCysValGlyGlyAspGlyMet 224
Db 522 ATGGCATCGGAGAGCTGGCGCGCTAGACGCTCTGGTGGTGCATGCTCGGAGACGGGTG 581
Qy 225 PheSerGluValLeuHisGlyLeuIleGlyValArgThrGlnArgSerAlaGlyValAspGln 244
Db 582 ATGCACGAGGTGGTGAACGGGCTCATGAGCGGCTGACTGGGAGACCGCCATCCAGAG 641
Qy 245 AsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAlaGly 264
Db 642 -----CCCTGTGTAGCTCCACGAGGC 665
Qy 265 SerThrAspCysValCys-----TyrSerThrValGlyThr 276
```

```
Db 666 TCTGCAACGCGTGGCAGCTTCCTTGAACCAATTATGCTGGCTATGAGCAGGTCAACAA 725
Qy 277 SerSepAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla--MetAsp 295
Db 726 GAAGACCTCTGCAACATGACGCTATTGTGTGCGCGCGCTGTCTACCCATGAAC 785
Qy 296 ValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyr 315
Db 786 CTGCTGCTCTGCACACGCGTTCGGGGCTGGCGCTCTCTCTGTCTGCTCAGCTGGCCCTGG 845
Qy 316 GlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArg 335
Db 846 GGCCTTCATTGCTGACACGCGTTCGGGGCTGGCGCTCTCTCTGTCTGCTCAGCTGGCCCTGG 905
Qy 336 TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSer 355
Db 906 TTCACTCTGGGCACCTTCCTCGCTGTGGCAGCCCTGTGGCAGCTACCGCGCGGAGCTGCC 965
Qy 356 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 375
Db 966 TACTCTCTCTAGGAGA--GTGGTTCCAGACACCTGCTCCGCC-----CTTGTG 1010
Qy 376 CysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyr 395
Db 1011 ---GTTGTGTCAGCAGGCGCGGTAGATGCACAC-----CTTGTG 1049
Qy 396 GlyLeuGluAlaAlaGluAspVal---GluGluTrpGlnValValCysGlyLys----- 412
Db 1050 CCACTGGAG-----GAGCCAGTCCCTCTCACTGGACAGTGGTGGCCGACGAGGACTTT 1103
Qy 413 -----PheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArg 430
Db 1104 GTGCTAGTCTCTGGCACTGCTGCATCGCACCTGGCAGTGAGATGTTGTGTCACCCATG 1163
Qy 431 GlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeuLeuAlaArgLys 450
Db 1164 GCGCGCTGTGCAGCT-----GGCGTCATGCATCTGTTCTACGTGCGGCGG 1208
Qy 451 ---CysSerArgPheAsnPheLeuArg---PheLeuIleArgHisThrAsnGlnGlnAsp 468
Db 1209 GGAGTGTCTCTGCGCATGCTGCTGCGCTCTTCCTGGCCATGSGAAGGCGGACGATATG 1268
Qy 469 GlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLys 488
Db 1269 GAGTATGATGCCCTACTTGGTATATGTGCGCGTGGTGGCTTCGCG----- 1316
Qy 489 HisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPheGlyHisIle 508
Db 1317 ---TTGGAGCCCAAGGAT-----GGAAAGGTGTGTTTGCA----- 1349
Qy 509 CysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsnCysAspGly 528
Db 1350 -----GTGGATGGG 1358
Qy 529 GluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPhe 548
Db 1359 GAATGTAGTGTAGGAGCGCGTGCAGGCGGCGAGGTGCACCCAACTACTTCTGGATGTC 1418
Qy 549 AlaArgGlyIleGluGluAsnPro-----LysPro 558
Db 1419 AGCGGTTGCGTGGAGCCCGCCGCGCAGCTGGAAGGCC 1454
```

RESULT 15

```
US-09-614-221A-399
; Sequence 399, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; WITH STEROL SYNTHESIS AND METABOLISM
```

```
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 399
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-399

Alignment Scores:
Pred. No.: 1.37e-10 Length: 1875
Score: 210.50 Matches: 66
Percent Similarity: 44.62% Conservative: 46
Best Local Similarity: 26.29% Mismatches: 102
Query Match: 6.96% Indels: 37
DB: 4 Gaps: 6
```

```
US-10-631-958-11 (1-562) x US-09-614-221A-399 (1-1875)

Qy 143 ThrLeuArgGluMetLeuLeuLysLeu-----ThrSerArgProLysHisLeuLeu 159
Db 631 ACTGTGGAAGAAATCTCTGGAAGAAAGTTATGAAATTTCCAAAGAGAAACAGATCGATATTA 690
Qy 160 ValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysVal 179
Db 691 GTCAATTATTATCCCAAGGTGTAAAGGTACTGCTAAAATTTATTTCTTGACAAAAGCA 750
Qy 180 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsn 199
Db 751 AGGCCAATACTAGTGGAAAGTGGCTGCAGAAATAGAAATTCATACACAAAAATATGCCCGT 810
Qy 200 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 219
Db 811 CACGCCCATCGATATTGCCAAGATTTAGATATCAGCAATACGATACCATTTGCGATGGCC 870
Qy 220 GlyGlyAspGlyMetPheSerGluValHisGlyLeuIleGlyArgThrGlnArgSer 239
Db 871 TCGGTGATGGTATTCATACGAAAGTAATTAATGGGCTTTATAGAAAGCCCGACAGA--- 927
Qy 240 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 259
Db 928 -----GTGGATGCGTTCAATAAACTAGCCGTAACTCAG----- 960
Qy 260 IleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAla 279
Db 961 ---CTACCTTTCGCGTTTCAGAAATTCATGAGCATTTTCATGTCAATGGCAAAATAACCCA 1017
Qy 280 GluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal----- 296
Db 1018 TCGTACCGCGCTCTGTCCTTGTCAATCCATTCGAAACAAGATAGACTTAATGTGTGT 1077
Qy 297 SerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGly 316
Db 1078 TCCCAACCTCTTATATGAACGAATGGCCAAGATTATCTCTTTTGTAGTCAGACAGCTACGGC 1137
Qy 317 PheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyr 336
Db 1138 GTTATTGCAGAAATCTGATATTAACACTATTCATCAGATGGATGGGTCCCGTTAGGT 1197
Qy 337 AspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPhe 356
Db 1198 AATTGGGT-----GTAGCATTCACATTATCCAAAGT----- 1230
Qy 357 LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 376
Db 1231 -----AAGAAATATCCCTCGCAAGTTTTCGTC 1257
Qy 377 PheValCysArgGlnSerLysGlnGlnLeuGlu 387
Db 1258 AAATATGCTGCCAAATCTAAAGGAATTAATA 1290
```

Search completed: September 6, 2005, 02:35:04
Job time : 367.875 secs



This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 5, 2005, 23:41:54 ; Search time 1593.07 Seconds
(without alignments)
2324.913 Million cell updates/sec

Title: US-10-631-958-11
Perfect score: 3025
Sequence: 1 HEANGPAPLGVRAAPPWRT.....QLVRLFARGIEENKPKDPSHS 562

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10631958/runat_02092005_165813_4470/app.query.fasta_1.1941
-DB=Published Applications NA -OPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10631958 @CGN 1 1 1588 @runat_02092005_165813_4470
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3025	100.0	1740	18	US-10-262-511-39	Sequence 39, Appl
2	3025	100.0	4413	10	US-09-969-896-16	Sequence 16, Appl
3	3025	100.0	4413	19	US-10-631-958-16	Sequence 16, Appl
4	3012	99.6	4429	19	US-10-618-941-55	Sequence 55, Appl
5	3004	99.3	4463	16	US-10-315-597A-1	Sequence 1, Appl
6	2931.5	96.9	4432	17	US-10-120-988-148	Sequence 148, App
7	2888	95.5	1614	10	US-09-969-896-9	Sequence 9, Appl
8	2888	95.5	1614	19	US-10-631-958-9	Sequence 9, Appl
9	2649.5	87.6	1840	9	US-09-784-810A-5	Sequence 5, Appl
10	2649.5	87.6	1840	22	US-10-876-281-5	Sequence 5, Appl
11	1640.5	54.2	979	10	US-09-969-896-1	Sequence 1, Appl
12	1640.5	54.2	979	19	US-10-631-958-1	Sequence 1, Appl
13	1156	38.2	817	19	US-10-115-635-247	Sequence 247, App
14	1008	33.3	550	10	US-09-969-896-8	Sequence 8, Appl
15	1008	33.3	550	19	US-10-631-958-8	Sequence 8, Appl
16	811	26.8	474	10	US-09-969-896-4	Sequence 4, Appl
17	811	26.8	474	19	US-10-631-958-4	Sequence 4, Appl
c 18	656	21.7	522	9	US-09-784-810A-7	Sequence 7, Appl
c 19	656	21.7	522	22	US-10-876-281-7	Sequence 7, Appl
20	616.5	20.4	2657	20	US-10-425-115-176413	Sequence 176413, A
21	573.5	19.0	2079	18	US-10-425-114-34530	Sequence 34530, A
22	573.5	19.0	2079	20	US-10-425-115-86211	Sequence 86211, A
23	525	17.4	329	10	US-09-969-896-5	Sequence 5, Appl
24	525	17.4	329	19	US-10-631-958-5	Sequence 5, Appl
c 25	480.5	15.9	2189	19	US-10-437-963-92801	Sequence 92801, A
26	437.5	14.5	564	22	US-10-477-445-46	Sequence 46, Appl
27	410	13.6	1833	19	US-10-437-963-30650	Sequence 30650, A
28	392.5	13.0	2165	20	US-10-425-115-64890	Sequence 64890, A
29	389	12.9	2084	18	US-10-424-599-47396	Sequence 47396, A
30	388	12.8	1869	18	US-10-425-114-34806	Sequence 34806, A
31	371	12.3	339	9	US-09-783-590-5271	Sequence 5271, Ap
32	363	12.0	3362	20	US-10-425-115-29002	Sequence 29002, A
33	347	11.5	2014	18	US-10-424-599-59419	Sequence 59419, A
34	344.5	11.4	1512	20	US-10-425-115-64892	Sequence 64892, A
35	344.5	11.4	1915	19	US-10-437-963-58701	Sequence 58701, A
36	343.5	11.4	382	9	US-09-784-810A-9	Sequence 9, Appl
37	343.5	11.4	382	22	US-10-876-281-9	Sequence 9, Appl
38	339	11.2	1774	18	US-10-425-114-34081	Sequence 34081, A
39	331.5	11.0	1386	19	US-10-437-963-95148	Sequence 95148, A
40	330.5	10.9	3008	19	US-10-437-963-77896	Sequence 77896, A
41	329.5	10.9	2609	17	US-10-348-052-25	Sequence 25, Appl
42	329.5	10.9	2609	19	US-10-622-011-25	Sequence 25, Appl
43	326.5	10.8	1857	18	US-10-425-114-32139	Sequence 32139, A
44	326.5	10.8	2247	20	US-10-425-115-97849	Sequence 97849, A
45	316.5	10.5	2284	20	US-10-425-115-79424	Sequence 79424, A

ALIGNMENTS

RESULT 1
US-10-262-511-39
; Sequence 39, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenn
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen

```
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Raestelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 39
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (76)..(1686)
US-10-262-511-39

Alignment Scores:
Pred. No.: 0 Length: 1740
Score: 3025.00 Matches: 562
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-631-958-11 (1-562) x US-10-262-511-39 (1-1740)

Qy 1 HisGluAlaAAsenGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThr 20
Db 1 CACGAGCGCGTAAACGGTCCGGCGCCCTCGCGCGTCCGCCGCCGCCAGCCCTGGCGGACG 60
Qy 21 SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrp 40
Db 61 AGCCCGCGCGGAGATGGGGCGGAGCGCGCGGCGGAGCGCGTGCATCCGTGCTGTGG 120
Qy 41 VallysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuLeuLeuTrp 60
121 GTGAAGCAGCAGCGCTCGCGCGTGCAGCTCGAGCCCGCGGGCTCTGCTGCGTGGTGG 180
61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAlaSpalaCysSerValProValSer 80
181 CGGAGCCCGGGCGCGGAGCGCGCCCGCGGTGCTGATGCTGCTCTGTGCGCTGTATCT 240
81 GluIleAlaValGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100
241 GAGATCATCGCGTTGAGGAAACAGAGCTTCACGGGAAACATCAAGGCAGTGGAAATGG 300
101 GlnLysMetGluLysProFyrAlaPheThrValHisCysValLysArgAlaArgHis 120
301 CAGAAATGGAAGACCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGCGCAC 360
121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrp 140
361 CGCTGGGAAGTGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGTGTGTCTTGTGG 420
141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal 160
421 CTGCAGACCTTGGCGGAGATGCTGGAGAAGCTGACGTCCAGACCACCAAGCATTTACTGGTA 480
161 PheIleAsnProPheGlyLysGlyGlnGlyLysArgIleFyrGluArgLysValAla 180
481 TTTATCAACCGTTTGGAGGAAAGGACAAGCGGATATATGAAGAAAAAGTGGCA 540
181 ProLeuPheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGln 200
541 CCACTGTTTCACTTAGCTTCATCCACTGACATCATCTGTTACTGAACATGCTTAATCAG 600
201 AlalysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 220
601 GCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCGCATCGTCTGTGCGGC 660
221 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla 240
661 GGAGATGATGTTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGAGCGCC 720
241 GlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 260
721 GGGGTCCACGACAGAACACCCCGGGCTGTGCTGGTCCCAGTAGCCTCCGAGTTTGGAAATC 780
261 IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu 280
781 ATTCCCGCAGSGTCAACGAGCTGCTGTGTTACTCCACCGTGGGCGCACGACGACGAGAA 840
281 ThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHis 300
841 ACCTCGCGCTGCATATCGTTGTTGGGAGCTCGCTGGGCCATGGATGTGTCTCAGTCCAC 900
301 HisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp 320
901 CACAACAGCACACTCTCTCCCTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACGGGAC 960
321 IleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGly 340
961 ATCATCAGGACAGTGGAGAGAACCGTGGTGGTCTTGCAGATACGACTTTTCAGGT 1020
341 LeuLysThrPheLeuSerHisHisCysTyrGlyThrValSerPheLeuProAlaGln 360
1021 TTAAGACCTTCTCTCCACCACTGCTATGAAGGAGCAGTGTCTCTCTCTCTCTCTCTCT 1080
361 HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg 380
1081 CACCGTGGGATCTCCAAGGATAGGAAGCCCTGCGCGGCGAGGATGCTTTGTTGAGG 1140
381 GlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAla 400
1141 CAAAGCAAGCAGCAGCTGGAGGAGGAGCAGACAGAAAGACACTGTATGTTTGGAAAGTGG 1200
401 GluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsn 420
```



```
Db 1201 GAGGACGTGGAGGAGTGGCAAGTCGTCTGTGGAGAGTTTCTGGCCATCAATGCCACCAAC 1260
Qy 421 MetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp 440
Db 1261 ATGTCCTGTGCTTGTGCGCGGAGCCCAAGGGGCTCTCCCGGCTGCCCACTTGGGAGAC 1320
Qy 441 GlySerSerAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 460
Db 1321 GGGTCTTCTGACCTCATCTCATCCGGAATGCTCCAGGTCAATTTTCTGAGATTTTCTC 1380
Qy 461 IleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValValArgVal 480
Db 1381 ATCAGGCACACCAACAGCAGGACAGTTTGACTTCACTTTTGTGAGTTTATCGCGTC 1440
Qy 481 LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly 500
Db 1441 AAGAAATTCAGTTTACGTCGAAGCACATGAGGATGAGGACAGCCACTCAAGGAGGGG 1500
Qy 501 GlyLysLysArgPheGlyHisLysCysSerSerHisProSerCysCysCysThrValSer 520
Db 1501 GGGAAAGAGCGCTTTGGGACATTTGACAGCAGCCACCCCTCTGCTGTCACCGTCTCC 1560
Qy 521 AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVal 540
Db 1561 AACAGCTCTGGACTGCGAGCGGGAGGTCTCGACACGCCCTGCCATCGAGGTTCAGAGTC 1620
Qy 541 HisCysGlnLeuValArgLeuPheAlaArgGlyLeuGluGluAsnProLysProAspSer 560
Db 1621 CACTGCCAGCTGGTTGCGACTCTTTGACAGAGGAATTGAAGAAATCCGAAGCCAGACTCA 1680
Qy 561 HisSer 562
Db 1681 CACAGC 1686
```

RESULT 2

```
; US-09-969-896-16
; Sequence 16, Application US/09969896
; Publication No. US2003012553A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
```

```
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
```

```
; LENGTH: 4413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-16
```

```
Alignment Scores:
Pred. No.: 0 Length: 4413
Score: 3025.00 Matches: 562
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
```

```
US-10-631-958-11 (1-562) x US-09-969-896-16 (1-4413)
```

```
Qy 1 HisGluAlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThr 20
Db 1 CACGAGGCGGCTAACGGTCCGCGCGCCCTCGCGCTCCGCGCGCCGCCAGCTGGCGGACG 60
Qy 21 SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrp 40
```

```
Db 61 AGCCCGCGCGGAGATGGGGCGACCGGGGCGCGGAGCCGCTGCAATCCGTGCTGTGG 120
Qy 41 ValLysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuLeuArgTrpTrp 60
Db 121 GTGAAGCAGCAGCAGCTCGCGCTGAGCGTGGAGCCCGCGCGGCTCTGCTGCGTGTGGT 180
Qy 61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAlaPalaCysSerValProValSer 80
Db 181 CGGAGCCCGCGGCGCGCGGAGCCCGCGCGCGGATGCGCTCTGTCGCCCTGTATCT 240
Qy 81 GluIleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100
Db 241 GAGATCATCGCCGTTGAGGAAACAGAGCTTCCGCGGAAACATCAAGCAGCTGGAAATGG 300
Qy 101 GlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHis 120
Db 301 CAGAAATGGAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGCAC 360
Qy 121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuLeuCysHisLeuTrp 140
Db 361 CGCTGGAGAGTGGCGCAGGTGACTTCTGTGGTGTCCAGAGGAGCAGCTGTGTCTGTGG 420
Qy 141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal 160
Db 421 CTGCAGACCCCTGCGGAGATGCTGGAGAGCTGACGTCAGACCCAGACCATTTACTGGTA 480
Qy 161 PheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleLysGluArgLysValAla 180
Db 481 TTTTATCAACCGTTTGGAGGAAAGGACAGGCAAGCGGATATATGAAAGAAAGTGGCA 540
Qy 181 ProLeuPheThrLeuAlaSerIleThrAspIleIleValThrGluHisAlaAsnGln 200
Db 541 CCACGTGTTCACTTTAGCCTCCATCACCACCTGACATCATCGTTACTGAAACATGCTAATCAG 600
Qy 201 AlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 220
Db 601 GCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACAGACGCGCATCGTCTGTGTCGGC 660
Qy 221 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla 240
Db 661 GGAGATGATGTTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGCGCC 720
Qy 241 GlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 260
Db 721 GGGGTCCAGCAGAACACCCCGCGGCTGTGCTGTCCTCCAGTAGCCTCCGATTGGAAATC 780
Qy 261 IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu 280
Db 781 ATTCGCCAGGGTCAACGGACTGCTGTGTGTACTCCACCGTGGGACGACGAGCGCAGAA 840
Qy 281 ThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHis 300
Db 841 ACCTCGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCCTCAGTCCAC 900
Qy 301 HisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp 320
Db 901 CACAACAGCACACTCTTCGCTACTCGCTGCTCCCTGCTGGGCTACGGCTCTCAGGGGAC 960
Qy 321 IleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaIleArgTyrAspPheSerGly 340
Db 961 ATCATCAAGGACAGTGAAGAAACGGTGGTGGGTCTTGGCAGATACGACTTTTCAGGT 1020
Qy 341 LeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGln 360
Db 1021 TTTAAAGACCTTCTCTCCACCACTGCTATGAAGGGGACAGTGTCTTCTCTCCCTGACAA 1080
Qy 361 HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg 380
Db 1081 CACACGCTGGGATCTCCAAGGGATAGGAAGCCCTGCGCGGCGAGGATGCTTTGTTGTCAGG 1140
Qy 381 GlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAla 400
Db 1141 CAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTCGG 1200
```



```
Db 1081 CACACGGTGGGATCTCAAGGAGTAGGAAGCCCTGCCGGCAGGATGCTTTGTTGCAGG 1140
Qy 381 GlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrGlyLeuAlaAla 400
Db 1141 CAAAGCAAGCAGCAGCTGGAGAGAGAGCAGAAAGAACACTGTATGGTTGGAACTGGC 1200
Qy 401 GluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsn 420
Db 1201 GAGGACGTGGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGCCCATCAATGCCACAAAC 1260
Qy 421 MetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp 440
Db 1261 ATGTCTGTGTGTGTGGCGAGGCCAGGGGCTCTCCCGGCTGCCACTTGGAGAC 1320
Qy 441 GlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeu 460
Db 1321 GGGTCTTCTGACCTCATCTCATCCGGAATGCTCCAGGTTTCAATTTTCTGAGATTTCTC 1380
Qy 461 IleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgVal 480
Db 1381 ATCAGGCACCAACACAGCAGGACAGTTTGACTTCTCACTTTTGTGAAGTTTATCGCGTC 1440
Qy 481 LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly 500
Db 1441 AAGAAATTCAGCTTACGTGGAAGCACATGGAGGATGAGCAGCGCTCAAGGAGGG 1500
Qy 501 GlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSer 520
Db 1501 CGGAAGAAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCTGCTGCTGCACCGCTCTCC 1560
Qy 521 AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVal 540
Db 1561 AACAGCTCTGGAACCTGCGACGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCAAGTTC 1620
Qy 541 HisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSer 560
Db 1621 CACTGCCAGCTGTTGCGACTCTTTGCACGAGGAATTGAAGAGAAATCCGAGCCACTCA 1680
Qy 561 HisSer 562
Db 1681 CACAGC 1686
```

RESULT 4

```
US-10-618-941-55
; Sequence 55, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 55
; LENGTH: 4429
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-618-941-55
```

Alignment Scores:

Pred. No.:	0	Length:	4429
Score:	3012.00	Matches:	560
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.57%	Indels:	0
DB:	19	Gaps:	0

US-10-631-958-11 (1-562) x US-10-618-941-55 (1-4429)

```
Qy 3 AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerPro 22
Db 23 GCCGCTAACCGTCCGGCGCCCTCGCGCTCCGCGCCGCCAGCGCTGGCGAGCGAGCCG 82
Qy 23 AlaAlaGluMetGlyAlaThrGlyValAlaGluProLeuGlnSerValLeuTrpValLys 42
Db 83 CGCGCGGAGATGGGGCGAGCGGGCGCGGAGCGCTGCAATCCGTCGCTGTGGGTGAAG 142
Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpArgSer 62
Db 143 CAGCAGCGCTGCCGCGTGAGCTGAGCCCGCGCGGGCTCTGCTGCGCTGTGGCGGAGC 202
Qy 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle 82
Db 203 CCGGGCGCGGAGCGCGCGCCCGCGCGGATGCTGCTCTGTGCTGTATCTGAGATC 262
Qy 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys 102
Db 263 ATCGCGCTTGAGGAAACAGACGTTTCA CGGGAAACATCAAGCAGTGGAAATGGCAGAA 322
Qy 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrp 122
Db 323 ATGGAAGAAGCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGCCGCTGG 382
Qy 123 LysTrpAlaGlnValThrPheTrpCysProGluGlnGlnLeuCysHisLeuTrpLeuGln 142
Db 383 AAGTGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGTGCAG 442
Qy 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162
Db 443 ACCCTGCGGAGATGCTGGAGAACTGACGTCCAGACCAAGCAATTTACTGGTATTATC 502
Qy 163 AsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu 182
Db 503 AACCGTTTGGAGGAAAGGACAAAGGCAAGCGGATATATGAAAGAAAAGTGGCACCCTG 562
Qy 183 PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys 202
Db 563 TTCACCTTAGCCTCCATCACCACCTGACATCATCTGTTACTGAACATGCTTAATCAGGCCAAG 622
Qy 203 GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysAsp 222
Db 623 GAGACTCTGTATGAGATTACATGACAAATACAGCGCATCGTCTGTGTGGCGGAGAT 682
Qy 223 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 242
Db 683 GGTATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGGTC 742
Qy 243 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro 262
Db 743 GACCAGAACCAACCCCGGGCTGTGTGTCCTCCAGTAGCTCCGGATTGGGAATCAATTCCC 802
Qy 263 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282
Db 803 GCAGGTCACCGACTGCGTGTGTGTACTCCACCGTGGGCACGAGCGCAGAAACCTTCG 862
Qy 283 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn 302
Db 863 CGCTGCATATCGTTTGTGGGACTCGCTGCGCCATGGATGTGCTCTCAGTCCACCACAAC 922
Qy 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322
Db 923 AGCACACTCTCTCGCTACTCGCTGTCTGCTGGGCTACGGCTTCTACGGGGACATCATC 982
Qy 323 LysAspSerGluLysValLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342
Db 983 AAGCAGTGTGAGAAAGAACCGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAG 1042
Qy 343 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 362
Db 1043 ACCTTCTCTCCACCACCTGCTATGAAGGGACAGTGTCTTCTTCTCCCTCCCTGCACACACG 1102
```



```
Db 1075 ACCTTCCTCCACACACTGCTATGAAGGACAGTGTCTTCTCCTCCCTGCACAAACACACAG 1134
QY 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382
Db 1135 GTGGGATCTCCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGCAGGCAAGC 1194
QY 383 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402
Db 1195 AAGCAGCAGCTGGAGAGAGAGAGAGAGAAAGCACTGTATGGTTTGAAGCTCGGAGGAC 1254
QY 403 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaAlaAsnAlaThrAsnMetSer 422
Db 1255 GTGGAGAGTGGCAAGTCTGTGTGGAGATTTCTGGCCATCAATGCCACAAACATGTCC 1314
QY 423 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442
Db 1315 TGTGCTTGTGCGCGGAGCCCAAGGGCCCTCTCCCGGCTGCCCACTTGGGAGCGGTCT 1374
QY 443 SerAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 462
Db 1375 TCTGACCTCATCTCATCCGGAATGCTCCAAAGTTCAATTTTCTGAGATTTCTCATCAGG 1434
QY 463 HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482
Db 1435 CACACCAACACAGACAGCAGTTTGACTTCACTTTTGTGAGTTTATCGCGTCAAGAAA 1494
QY 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys 502
Db 1495 TTTCCAGTTTACCTCGAAGCACATGAGGATGAGGACAGCAGCTCAAGGAGGGGGGAAG 1554
QY 503 LysArgPheGlyHisLysCysSerSerHisProSerCysCysCysThrValSerAsnSer 522
Db 1555 AAGCCCTTTGGGCACATTTGACAGCAGCCACCCCTCTGCTGTCACCGTCTTCCAAACAGC 1614
QY 523 SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys 542
Db 1615 TCTTGGAACTGCGATGGGAGGTCCTGCACACCCCTGCCATCGAGTTCAGTCCACTGC 1674
QY 543 GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 562
Db 1675 CAGCTGGTTCGACTCTTTGCACGAGGAATTTGAAGAGATCCGAAGCCAGACTCACACAGC 1734
```

RESULT 6

```
US-10-120-988-148
; Sequence 148, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyun
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 148
; LENGTH: 4432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129) .. (2817)
```

```
US-10-120-988-148
Alignment Scores: 4.03e-313 Length: 4432
Pred. No.:
```

```
Score: 2931.50 Matches: 548
Percent Similarity: 98.21% Conservative: 2
Best Local Similarity: 97.86% Mismatches: 7
Query Match: 96.91% Indels: 3
DB: 17 Gaps: 1

US-10-631-958-11 (1-562) x US-10-120-988-148 (1-4432)
QY 6 GlyProAlaProLeuGly-----ValArgAlaProProAlaTrpArgThrSerPro 22
Db 1135 GGTTCGCCCTATAGACAAGACAGTCAAGAGAGCTGCAGGCTTGGCGGACGAGCCCG 1194
QY 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys 42
Db 1195 GCGGCGAGATGGGGGCGACCGGGCGCGGAGCCGCTGCAATCCGTGCTGTGGGTGAAG 1254
QY 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer 62
Db 1255 CAGCAGCGCTGGCGCGTGAGCTGGAGCCCGCGGGGCTCTGCTGCGTGTGTGGCGGAGC 1314
QY 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle 82
Db 1315 CCGGGCCCGGAGCGCGGCCCGCGCGATGCTCTCTGTGCTGTATCTGAGATC 1374
QY 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys 102
Db 1375 ATCCCGCTTGAGGAAACAGACGTTTACGGGAAACATCAAGGCAGTGGGAAAATGGCAGAAA 1434
QY 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrp 122
Db 1435 ATGGAAAAGCCTTACGCTTTTACAGTTTCTGTGTAAAGAGACGACGCGGACCGCTGG 1494
QY 123 LysTrpAlaGlnValThrPheTrpCysProGluGlnGlnLeuCysHisLeuTrpLeuGln 142
Db 1495 AAGTGGCGCAGGTGACTTTCTGTGTCTCCAGAGGAGCAGTGTGTCTGTGGCTGCAG 1554
QY 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162
Db 1555 ACCCTGCGGGAGATGCTGGAGAAGCTGACGTCCAGACCAAAAGCATTTTACTGGTATTATC 1614
QY 163 AsnProPheGlyLysGlyGlnGlyLysArgIleTyrGluAtgLysValAlaProLeu 182
Db 1615 AACCCGTTTGGAGGAAAAGGACAGGCAAGCGGATATATGAAGAAAAGTTGGCACCATG 1674
QY 183 PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys 202
Db 1675 TTCACCTTAGCCTCCATCACCACATGACATCATCGTTACTGAACATGCTAATCAGGCCAAG 1734
QY 203 GluThrLeuTyrGluIleAsnIleAspLysTrpAspGlyIleValCysValGlyGlyAsp 222
Db 1735 GAGACTCTGTATGAGATTAACATAGACAAATACGACGCGCATCGTCTGTGTGCGCGGAGAT 1794
QY 223 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 242
Db 1795 GGTATGTTTCAGCGAGGTGCTGACGCTGTGATTGGAGAGACGACGAGGAGCCCGGGTTC 1854
QY 243 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro 262
Db 1855 GACCAGAACACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGATTTGGAATCATTCCTCC 1914
QY 263 AlaGlySerThrAspCysValCysTrpSerThrValGlyThrSerAspAlaGluThrSer 282
Db 1915 GCAGGGTCAACCGACTGCGTGTGTACTTCCACCGCTGGGCGACGACGACGAGAAACCTCG 1974
QY 283 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsn 302
Db 1975 GCGCTGCATATCGTTGTTGGGAGCTCGCTGGCCATGATGTCTCTAGTCCACCAAC 2034
QY 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322
Db 2035 AGCACACTCTTCTGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2094
QY 323 LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342
```

```
Db 2095 AAGGACAGTGAGAAACCGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAG 2154
Qy 343 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 362
Db 2155 ACCTTCTCTCTCCACCACTGCTATGGAGGGACAGTGTCTTCTCCCTGCAACAACACG 2214
Qy 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382
Db 2215 GTGGGATCTCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTGACGGCAAGC 2274
Qy 383 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402
Db 2275 AAGCAGCAGCTGGAGGAGCAGAGAAACACTGTATGTGTTTGGAAAGCTGCGGAGAC 2334
Qy 403 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAenAlaThrAsnMetSer 422
Db 2335 GTGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAAATGTGTC 2394
Qy 423 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442
Db 2395 TGTGCTTGTCCCGGAGCCCGCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCT 2454
Qy 443 SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg 462
Db 2455 TCTGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG 2514
Qy 463 HisThrAsnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482
Db 2515 CACACCAACCGAGGACCACTTTGACTTCACCTTTTGTGAAGTTTATCGCGTCAAGAAA 2574
Qy 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys 502
Db 2575 TTCCAGTTACGTCGAGCACATGGAGGATGAGGACAGGACCTCAAGAGGGGGGGAAG 2634
Qy 503 LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSer 522
Db 2635 AAGCGCTTGGGCACATTTGCAGCAGCCACCCCTCTGCTGCTGCACCGTCTCCAACAGC 2694
Qy 523 SerTrpAsnCysAspGlyValLeuHisSerProAlaIleGluValAtqValHisCys 542
Db 2695 TCCTGGAACTGCGACGGGGAGGTCTGCACAGCCCTGCCATCGAGGTCAAGAGTCCACTGC 2754
Qy 543 GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 562
Db 2755 CAGCTGGTTCGACTCTTTCACGAGGAATTGAAGAAATCCGAAGCCAGACTCACACAGC 2814
```

RESULT 7

```
US-09-969-896-9
; Sequence 9, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-9
```

```
Alignment Scores: 6.05e-309 Length: 1614
Pred. No.: 2888.00 Matches: 537
Score:
```

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.47% Indels: 0
DB: 10 Gaps: 0
US-10-631-958-11 (1-562) x US-09-969-896-9 (1-1614)
Qy 26 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 45
Db 1 ATGGGGCGGACGGGGCGGGCGGAGCCGCTGCAATCCGTCGTGTGGTGAAGCAGCAGCGC 60
Qy 46 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 65
Db 61 TGCSCCGTGAAGCTGGAGCCCGCGGGGCTCTGCTGGCTGGTGGCGAGCCCGGGGCC 120
Qy 66 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 85
Db 121 GGAGCCGGCGCCCGCGGGGATGCTCTCTGTGCTGTATCTGAGATCATCGCGTT 180
Qy 86 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 105
Db 181 GAGGAAACAGACGTTTCAAGGGAAACATCAAGGCAGTGGAAAAATGGCAGAAAAATGGA 240
Qy 106 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 125
Db 241 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGGCACCCGCTGGAAGTGGCG 300
Qy 126 GlnValThrPheTrpCysProGluGlnLysCysHisLeuTrpLeuGlnThrLeuArg 145
Db 301 CAGGTGACTTTTCTGGTGTCCAGAGGAGCAGCTGTGTCTGTGTGGCTGACACCTCGCG 360
Qy 146 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 165
Db 361 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGTGTTATTTATCAACCCGTT 420
Qy 166 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 185
Db 421 GGAGGAAAGACAAAGCAAGCGGATATATCAAAAGAAAAGTGGCACCACCTGTTTACCTTA 480
Qy 186 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 205
Db 481 GCCTCCATACCACTGACATCATCTGTTACTGAAACATGCTTAATACAGGCCAAGGACACTG 540
Qy 206 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 225
Db 541 TATCAGATTAACTAGACAAATACGACGGCATCTCTGTGTCCGGGAGATGGTATGTTTC 600
Qy 226 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 245
Db 601 AGCGAGGTGCTGCACGGTCTGATTTGGGAGGACGCGAGGAGCGCGGGGTCCAGCAGAAC 660
Qy 246 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 265
Db 661 CACCCCGGGCTGTCTGGTCCCCAGTAGCTCCCGAATCGGAATCATTTCCCGCAGGGTCA 720
Qy 266 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 285
Db 721 ACGGACTGCGTGTGTTACTCCACCGTGGGCACCGACGACGAGAAACCTCGCGCTGCAT 780
Qy 286 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 305
Db 781 ATCGTTGTTGGGACTCGCTGGCCATGGATGTGCTCTCAGTCCACCAACAGCAGCACTC 840
Qy 306 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 325
Db 841 CTTCCGTACTCCGTCCTGCTGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT 900
Qy 326 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 345
Db 901 GAGAAAGAAACGTTGTTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTC 960
Qy 346 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 365
Db 1614
Matches: 537
```

Db 961 TCCACCACCTGCTATGAGGAGCAGTGCTCTTCTCCCTGTCACACACACAGCGTGGGATCT 1020
QY 366 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 385
Db 1021 CCAAGGGATAGGAAGCCCTGCCGGGCGAGCATGCTTTGTTGCGAGCAAGCAGCAG 1080
QY 386 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 405
Db 1081 CTGGAGGAGGAGCAGAAAGACACTGTATGTGTTGGAAGCTGCGAGGACGCTGGAGGAG 1140
QY 406 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 425
Db 1141 TGGCAAGTCGTCTGTGGGAAGTTCTGGCCCATCAATGCCACAAACATGTCCTGTCTGT 1200
QY 426 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 445
Db 1201 CGCCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1260
QY 446 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 465
Db 1261 ATCCCTCATCCGGAATGCTCCAGGTTCAAATTTCTGAGATTCTCATCAGGCACACCAAC 1320
QY 466 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 485
Db 1321 CAGCAGGACCACTTTGACTTCTACTTTGTTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
QY 486 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 505
Db 1381 ACGTCGAAGCACATGAGGATGAGGACAGCAGCCTCAAGAGGGGGGGAAGAGCGCTTT 1440
QY 506 GlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsn 525
Db 1441 GGGCACATTTGACGAGCAGCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
QY 526 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 545
Db 1501 TCGCAGGGGAGGTCTCGACAGCCCTGCCATCGAGGTGAGAGTCCAGTCCAGCTGGTT 1560
QY 546 ArgLeuPheAlaArgGlyIleGluAsnProLysProAspSerHisSer 562
Db 1561 CGACTCTTTGACGAGGAATTTGAAGAAATCCGAAGCAGACTCACACAGC 1611

RESULT 8

US-10-631-958-9
; Sequence 9, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIORITY FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-9

Alignment Scores:

Pred. No.: 6.05e-309 Length: 1614
Score: 2888.00 Matches: 537
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.47% Indels: 0

DB: 19 Gaps: 0
US-10-631-958-11 (1-562) x US-10-631-958-9 (1-1614)
QY 26 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 45
Db 1 ATGGGGGGGCGAGCGGGGCGGAGCGCTGCATTCCTGCTGTGGGTGNAGCAGCAGCGC 60
QY 46 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgGlnTrpTrpArgSerProGlyPro 65
Db 61 TGGCGGTGAGCTGGAGCGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 66 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 85
Db 121 GGAGCGGCGGCGCGCGCGGCGGATGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 86 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 105
Db 181 GAGGNAACAGACGTTACGGGNAACATCAAGCAGTGGNAATGCGAGNAATGCGAANAAG 240
QY 106 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgHisArgTrpLysTrpAla 125
Db 241 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 300
QY 126 GlnValThrPheTrpCysProGluGluGlnCysHisLeuTrpLeuTrpLeuTrpLeuArg 145
Db 301 CAGGTGACTTTCTGGTGTCCAGAGGAGCAGTGTGTCTACTTGTGTGTGTGTGTGTGTGTGTGT 360
QY 146 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 165
Db 361 GAGATGCTGGAGAGTGTACGTCCAGACCAAGCATTTTACTGTGTATTTATCAACCCGTTT 420
QY 166 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 185
Db 421 GGAGAAAGACACAAGGCAAGCGGATATGAAAGAAAGTGGCACCCTGTTTCACTTCA 480
QY 186 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 205
Db 481 GCCTCCATCACCCTGACATCATCTGTTACTGAACATGCTTAATCAGCCCAAGGAGACTCTG 540
QY 206 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysAspGlyMetPhe 225
Db 541 TATGAGATTACATAGACAAATACAGCGCATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 226 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 245
Db 601 AGCGAGGTGCTGCAGCGTCTGATTGGGAGGAGCAGCAGCAGCGCGCGGTGCGACAGAAC 660
QY 246 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProIleArgSer 265
Db 661 CACCCCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY 266 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 285
Db 721 ACGGACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 286 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 305
Db 781 ATCGTTTGGGACTCGCTGGCCATGATGTCTCTCAGTCCACCAACCAACAGCACATC 840
QY 306 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 325
Db 841 CTTGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 326 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 345
Db 901 GAGAGAAACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
QY 346 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 365
Db 961 TCCCACTGCTATGAGGGGACAGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
QY 366 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 385

Ds	1021	CCAAAGGATAGAAAGCCCTGCGGCGCAGATGCTTTGTTTGCAGGCAAGCAAGCAGCAG	1080	
Qy	386	LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaGluAAspValGluGlu	405	
Ds	1081	CTGAGGAGGAGCAGAGAAAGCACTGATGTTGGTGGTGGAGGAGCTGGAGGAG	1140	
Qy	406	TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys	425	
Ds	1141	TGGCAAGTCGTCGTGGGAAGTTCTGGCCATCAATGCCACAACATGCTCTGTGCTGT	1200	
Qy	426	ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu	445	
Ds	1201	CGCCGGAGCCCAAGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	1260	
Qy	446	IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn	465	
Ds	1261	ATCCTCATCCGGAATGCTCCAGGTTCAAATTTCTGAGATTTCTCATCAGGACACCAAC	1320	
Qy	466	GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe	485	
Ds	1321	CAGCAGGACCAAGTTGACTTCACTTTGTGGAAGTTTATCGGTCAGAAATTCAGTTT	1380	
Qy	486	ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe	505	
Ds	1381	ACGTCGAAGCACATGGAGGATGAGACAGCGACCTCAAGGAGGGGGGGAAGCGCTTT	1440	
Qy	506	GlyHisLysCysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsn	525	
Ds	1441	GGGCACATTTGAGCAGGACCCCTCTGCTGCTGTCACCGTCTCAACAGCTCTCTGGAAC	1500	
Qy	526	CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal	545	
Ds	1501	TGCGACGGGGAGGCTCTGCACAGCCCTGCCATCGAGGTCAGAGTCCAGCTGGCTT	1560	
Qy	546	ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer	562	
Ds	1561	CGACTCTTTGCAGCAGGAATTGAAGAGAATCCGAAGCCAGACTCACACAGC	1611	
RESULT 9				
US-09-784-810A-5				
; Sequence 5, Application US/09784810A				
; Patent No. US20020082203A1				
; GENERAL INFORMATION:				
; APPLICANT: RASTELLI, LUCA				
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING				
; TITLE OF INVENTION: SAME				
; FILE REFERENCE: 10716-08				
; CURRENT APPLICATION NUMBER: US/09/784,810A				
; CURRENT FILING DATE: 2001-02-14				
; PRIOR APPLICATION NUMBER: 60/182,360				
; PRIOR FILING DATE: 2000-02-14				
; PRIOR APPLICATION NUMBER: 60/191,261				
; PRIOR FILING DATE: 2000-03-22				
; NUMBER OF SEQ ID NOS: 29				
; SOFTWARE: PatentIn Ver. 2.1				
; SEQ ID NO 5				
; LENGTH: 1840				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
US-09-784-810A-5				
Alignment Scores:				
Pred. No.:	1,9e-282	Length:	1840	
Score:	2649.50	Matches:	495	
Percent Similarity:	97.83%	Conservative:	1	
Best Local Similarity:	97.63%	Mismatches:	0	
Query Match:	87.59%	Indels:	11	
Ds:	9	Gaps:	1	
US-10-631-958-11 (1-562) x US-09-784-810A-5 (1-1840)				
Qy	67	AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaValGlu	86	

Ds	1	GCCGCGCCCGCGCGCGGATGCCCTGCTCTGTGCTCTATCTGAGATCATCGCGTTGAG	60	
Qy	87	GluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro	106	
Ds	61	GAACACAGCGTTACGGGAACATCAAGGCGAGTGGAAAAATGGCAGAAATATGAAAAAGCCT	120	
Qy	107	TyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAlaGln	126	
Ds	121	TACGCTTTTACAGTTCACTGTGTAAGAGAGACGACGCGCACCGCTGGAAGTGGCGCAG	180	
Qy	127	ValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu	146	
Ds	181	GTGACTTTTCTGGTCTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCGAGACCTCGCGGAG	240	
Qy	147	MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly	166	
Ds	241	ATGCTGGAGAGCTGAGCTCCAGACCAAGCAATTTACTTGGTATTTATCAACCCCGTTTGA	300	
Qy	167	GlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAla	186	
Ds	301	GGAAAAGACACAGCAAGCGGATATATGAAAGAAAGTGGCACCACTGTTCCACCTTAGCC	360	
Qy	187	SerIleThrThrAspIleIle-----ValThr	195	
Ds	361	TCCATCACCACATGACATCATCGGTAAACAAATTCCTATGTTAACTATGTAGAAGTAATTA	420	
Qy	196	GluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGly	215	
Ds	421	GAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAATACGACGGC	480	
Qy	216	IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArg	235	
Ds	481	ATCGTCTGTGTCGGCGAGATGGTATGTTTCAGCGAGGTGCTGCACCGTCTGATTGGGAGG	540	
Qy	236	ThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSer	255	
Ds	541	ACGCAGAGGACGCGCGGGGTGACACAGAACACCCCGGGGTGTGCTGGTCCCAGTAGC	600	
Qy	256	LeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGly	275	
Ds	601	CTCCGGATTGGATCATTTCCGCGAGGTCAACGAGCTGCGTGTGTTACTCCACCGTGGGC	660	
Qy	276	ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAsp	295	
Ds	661	ACCAGCAGCAGAAAACCTCGGCGCTGCATATCGTTGTGGGACTCGCTGGCCATGGAT	720	
Qy	296	ValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyr	315	
Ds	721	GTGTCTCTCAGTCCACCAACACAGCACACTCCCTCGCTACTCCGTGTCCTGCTGGGCTAC	780	
Qy	316	GlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArg	335	
Ds	781	GGCTTCTACGGGGACATCATCAAGACAGTGAAGAAACCGTGGTGGTCTTGCCAGA	840	
Qy	336	TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSer	355	
Ds	841	TACGACTTTTACGGTTTAAAGACCTTCTCTCTCCACCACTGCTATGAAGGACAGTGTCC	900	
Qy	356	PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly	375	
Ds	901	TTCTCTCCCTGCACAAACACACCGTGGATCTCTCAAGGGATAGGAAGCCCTGCGCGGACGA	960	
Qy	376	CysPheValCysArgGlnSerLysGlnLeuGluGlnLysLysLysAlaLeuTyr	395	
Ds	961	TGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAAGAAAGCACTGTAT	1020	
Qy	396	GlyLeuGluAlaAlaGluAspValGluTrpGlnValValCysGlyLysPheLeuAla	415	
Ds	1021	GGTTTGAAGCTGCGGAGGACGTGGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCC	1080	
Qy	416	IleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAla	435	


```
Db 1201 TTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCAGTGTGACTTTCACCTTTGGTT 1260
Qy 476 GluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSer 495
Db 1261 GAAGTTTATCGCGTCAAGAAATTCAGTTTACGTTCGAAGCAATGGAGGATGAGGACAGC 1320
Qy 496 AspLeuLysGluGlyLysLysArgPheGlyHisLysCysSerSerHisProSerCys 515
Db 1321 GACCTCAAGGAGGGGGGAAGAGCGCTTTGGGCACATTTGCAGACGCCCTCTCTGC 1380
Qy 516 CysCysThrValSerAsnSerSerTriaAsnCysAspGlyGluValLeuHisSerProAla 535
Db 1381 TGCTGCACCGTCTCCAAACAGCTCTCGAATGCGACGGGAGGTCTCTGCACAGCCCTGCC 1440
Qy 536 IleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyLysGluAsn 555
Db 1441 ATCAGAGTCAGGGTCCACTGCCAGCTGGTTCGACTCTTTTGACGAGGAATTGAAGAAAT 1500
Qy 556 ProLysProAspSerHisSer 562
Db 1501 CCGAAGCCAGACTCACACAGC 1521

RESULT 11
US-09-969-896-1
; Sequence 1, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR FILING DATE: 2000-10-06
; PRIOR FILING DATE: 2000-10-06
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-1

Alignment Scores:
Pred. No.: 4.17e-171 Length: 979
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 54.23% Indels: 11
DB: 1.0 Gaps: 1

US-10-631-958-11 (1-562) x US-09-969-896-1 (1-979)
Qy 155 ProLysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 174
Db 2 CCAAGCAATTACTGGTATTATCAACCGTTTGAGGAAAGGACAGGCGGAGGATA 61
Qy 175 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle--- 193
Db 62 TATGAAGAAAGTGGCACCACCTGTTACCTTAGCTTACCTCCTCATCACCCTGACATCATCGGT 121
Qy 194 -----ValThrGluHisAlaAsnGlnAlaLysGlu 203
Db 122 AACAAATTTCTATGTTAACTATGTAGAAGTAATTACTGAACATGCTAATCAGGCCAAGGAG 181
Qy 204 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyLysValGlyLysValGlyLysGly 223
Db 182 ACTCTGTATGAGATTAAACATAGCAATACACGCGCATCGTGTGTGCGGAGATGGT 241
Qy 224 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 243
```

```
Db 242 ATGTTTACGAGGCTGCTGCACCGTCTGATTGGGAGGACGACGAGGAGCGCGGGCTCGAC 301
Qy 244 GluAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyLysIleProAla 263
Db 302 CAGAACCACCCCGGGCTGTGGTCTCCAGTAGCTCCGGATTTGGAATCAITTTCCCGCA 361
Qy 264 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 283
Db 362 GGGTCAACGAGCTGCGGTGTGTACTCCACCGTGGGACACGACGACGACGAGAACCTCGCGC 421
Qy 284 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 303
Db 422 CTGCATATCGTTGTGGGACTCGCTGGCCATGATGTCTCAGTCCAGTCCACCAACACAGC 481
Qy 304 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 323
Db 482 ACACTCTCTCGCTACTCCGTGTCCTGCTGGGTACGGCTTCTACGGGGACATCATCAAG 541
Qy 324 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 343
Db 542 GACAGTGAGAAGAAACGCTGGTGGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACC 601
Qy 344 PheLeuSerHisHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThrVal 363
Db 602 TTCCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACGCTG 661
Qy 364 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 383
Db 662 GSATCTCCAAAGGATAGGAAGCCCTCCGGGAGGAGTGTGTTGTTTCAGGCAAGCAAG 721
Qy 384 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 403
Db 722 CAGCAGCTGGAGGAGGAGCAGAGAAGCACTGTATGTTTGGAGCTGCGGAGGACGTG 781
Qy 404 GluGlnTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 423
Db 782 GAGGAGTGCCAAAGTCGTCTGTGGAAAGTTTCTGCCATCAATGCCACAAACATGCTCTGT 841
Qy 424 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 443
Db 842 GCTTGTGCGCGAGCCCGGCGGCTCTCCCGGCTGCCCATTTCTGAGATTTCATCAGGCAC 901
Qy 444 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 463
Db 902 GACCTCATCTCATCCGGAATGCTCCAGGTTCATTTTCTGAGATTTCATCAGGCAC 961
Qy 464 ThrAsnGlnGlnAspGln 469
Db 962 ACCAACCCAGCAGGACCAG 979

RESULT 12
US-10-631-958-1
; Sequence 1, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
```

; ORGANISM: Homo sapiens			
US-10-631-958-1			
Alignment Scores:			
Pred. No.:	4.17e-171	Length:	979
Score:	1640.50	Matches:	314
Percent Similarity:	96.63%	Conservative:	1
Best Local Similarity:	96.32%	Mismatches:	0
Query Match:	54.23%	Indels:	11
DB:	19	Gaps:	1
US-10-631-958-11 (1-562) x US-10-631-958-1 (1-979)			
Qy	155	ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle	174
Db	2	CCAAAGCATTTACTGGTATTTATCAACCCGTTTGAGGAAAGGCAAGGCAAGCGGATA	61
Qy	175	TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle---	193
Db	62	TATGAAGAAGATGCGACCACTGTTACCTTAGCTTCATCCACCATGACATCATCGGT	121
Qy	194	-----ValThrGluHisAlaAsnGlnAlaLysGlu	203
Db	122	AACAAATTTCTATGTTAACTATGTAGAGTAATTAATCTGAACATGCTAATCAGGCCAAGCAG	181
Qy	204	ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly	223
Db	182	ACTCTGTATGAGATTAAATACAAATACGACGCGCATCGTGTGTGCGGGAGATGGT	241
Qy	224	MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp	243
Db	242	ATGTTCCAGGAGTCTGACCGTCTGATGGAGAGCGAGAGGCGCGGGTGCAC	301
Qy	244	GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla	263
Db	302	CAGAACCAACCCCGCGCTGTGCTGCTCCCGATGAGTCCCGATTTGAATCATTTCCGCA	361
Qy	264	GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla	283
Db	362	GGGTCAACCGGACTGGCTGTGTACTCCACCGTGGGCACGACGACGAGAACCTCGGCG	421
Qy	284	LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer	303
Db	422	CTGCATATCGTTTGGGAGCTCGCTGGCCATGGATGTCTCTAGTCCACCAACACAGC	481
Qy	304	ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys	323
Db	482	ACACTCCTTCGCTACTCCGCTGTCCTGCTGGGCTACGGCTTCTAGGGGGACATCATCAAG	541
Qy	324	AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr	343
Db	542	GACAGTGAGAGAAACCGGTGGTGTGGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACC	601
Qy	344	PheLeuSerHisHisCysTyrGluClyThrValSerPheLeuProAlaGlnHisThrVal	363
Db	602	TTCTCTCTCCCACTCTGCTATGAAGGAGACGTGCTCTTCTCCCTGCAACAACACCGGTG	661
Qy	364	GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys	383
Db	662	GGATCTCCAGGATAGGAGCCCTGCCGGCAGGATGCTTTGTTTTCAGGCNAAGCAG	721
Qy	384	GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal	403
Db	722	CAGCAGCTGGAGGAGGAGCAGAAAGCACTGTATGGTTTGGAGCTGCGGAGCAGCTG	781
Qy	404	GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys	423
Db	782	GAGGAGTGGCAAGTCTGTGGGGAAGTTTCTGGCCATCAATGCCACAAAATGTCCTGT	841
Qy	424	AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer	443
Db	842	GCTTGTCCCGGAGGCCCGAGGGGCTCTCCCGGGCTGCCACTTGGGAGACGGGTCTTCT	901

Qy	444	AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis	463
Db	902	GACCTCATCTCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC	961
Qy	464	ThrAsnGlnGlnAspGln	469
Db	962	ACCAACCCAGCAGGACCAG	979
RESULT 13			
US-10-115-635-247			
; Sequence 247, Application US/10115635			
; Publication No. US20040137434A1			
; GENERAL INFORMATION:			
; APPLICANT: Tang, Y. Tom			
; APPLICANT: Zhou, Ping			
; APPLICANT: Goodrich, Ryle			
; APPLICANT: Liu, Chenghua			
; APPLICANT: Asundi, Vinod			
; APPLICANT: Ren, Feiyan			
; APPLICANT: Zhang, Jie			
; APPLICANT: Zhao, Qing A.			
; APPLICANT: Xue, Aidong J.			
; APPLICANT: Yang, Yonghong			
; APPLICANT: Wehrman, Tom			
; APPLICANT: Drmanac, Radoje T.			
; TITLE OF INVENTION: Novel Nucleic Acids and			
; FILE REFERENCE: Polypeptides			
; CURRENT APPLICATION NUMBER: US/10/115,635			
; CURRENT FILING DATE: 2002-04-03			
; PRIOR APPLICATION NUMBER: 09/714,936			
; PRIOR FILING DATE: 2000-11-17			
; NUMBER OF SEQ ID NOS: 362			
; SOFTWARE: pt FL_genes Version 2.0			
; SEQ ID NO 247			
; LENGTH: 817			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (90)..(662)			
US-10-115-635-247			
Alignment Scores:			
Pred. No.:	1.45e-117	Length:	817
Score:	1156.00	Matches:	213
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	38.21%	Indels:	0
DB:	19	Gaps:	0
US-10-631-958-11 (1-562) x US-10-115-635-247 (1-817)			
Qy	3	AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProAlaTyrArgThrSerPro	22
Db	21	GCCGCTAACGGTCGCGCGCCCTCGCGCTCGCGCGCCCGCCAGCTCGCGACGAGCCCG	80
Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTyrValLys	42
Db	81	CGCGCGAGATGGGGGCGACGGGGGCGCGGAGCCGTCGCATCCGTGCTGGGTGAG	140
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTyrArgSer	62
Db	141	CAGCAGCGCTGCGCGGTGAGCTGGAGCCCGCGCGGCTCTGCTGCGCTGGTGGGAGC	200
Qy	63	ProGlyProGlyAlaGlyAlaProGlyAlaAsnAlaCysSerValProValSerGluIle	82
Db	201	CCGGGGCCGAGCGCGCGCCCGCGCGGATGCCCTCTGTGCTGTATCTGAGATC	260
Qy	83	IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTyrGlnLys	102
Db	261	ATCGCCGTTGAGGNAACAGAGCTTCACGGGNAACATCAAGGCAGTGGAAAATGCGAGNA	320

Db	181	CGGAGCCCGGGCCCGGAGCGGGCCCCCGGCGGATGCTGCTGTGCTGTATCT	240
Qy	81	GluIleIleAlaValGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp	100
Db	241	GAGATCATCGCGGTTGAGGAAACAGACGTTCAACGGGAAACATCAAGGCAGTGGAAAAATGG	300
Qy	101	GlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHis	120
Db	301	CAGAAATGGAAAGCCTTTACAGTTTCACTGTGTAAAGAGACGACGGCAC	360
Qy	121	ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrp	140
Db	361	CGCTGGAAGTGGGCGCAGGTGACTTCTGGTGTCCAGAGGACGCTGTGTCACTTGTGG	420
Qy	141	LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuVal	160
Db	421	CTGCAGACCCCTCGGGGAGATGCTGGAGAAGCTGACGTCAGACCAAGCATTTACTGGTA	480
Qy	161	PheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAla	180
Db	481	TTTATCAACCCGTTTGGAGGAAAGGACAAAGCCGCGATATATGAAAGAAAAAGTGGCA	540
Qy	181	ProLeuPhe	183
Db	541	CCACTGTTC	549

Search completed: September 6, 2005, 10:49:17
Job time : 1646.07 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 6, 2005, 02:19:54 ; Search time 867.649 Seconds
(without alignments)
3834.378 Million cell updates/sec

Title: US-10-631-958-11

Perfect score: 3025

Sequence: 1 HEANGPAPLGVAPPWRT.....QLVRLFARGIENKPDPSHS 562

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10631958/runat.02092005.165814.4497/app.query.fasta_1.1941
-DB=N Geneseq 16Dec04 -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CGN.1.1436 @runat.02092005.165814.4497 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3025	100.0	1740	8	ADA05679
2	3025	100.0	4413	6	ABL40828
3	3012	99.6	4429	12	ADJ96598
4	3012	99.6	4445	13	ADP55247
5	3004	99.3	4463	6	ABA96945

6	2980	98.5	1740	12	ADNG2844
7	2931.5	96.9	4432	8	ABX70921
8	2649.5	87.6	1840	5	AAAD14426
9	2587	85.5	4231	3	AAA50510
10	2541.5	84.0	4702	13	ADS10370
11	2400	79.3	2241	5	AA577728
12	2210	73.1	3975	8	ACC70838
13	1714.5	56.7	2186	5	AA577730
14	1640.5	54.2	979	6	ABL40822
15	1386.5	45.8	1570	5	AA577731
16	1156	38.2	817	6	ABQ99499
17	1156	38.2	817	13	ADS11669
18	1055	34.9	753	3	AA576031
19	1008	33.3	550	6	ABL40827
20	986	32.6	547	3	AA576592
21	811	26.8	474	6	ABL40823
22	721.5	23.9	10337	4	ABL25704
23	686.5	22.7	2173	4	ABL25705
24	679	22.4	426	5	AA577727
25	656	21.7	522	5	AAAD14427
26	626	20.7	411	5	AA577729
27	601	19.9	1774	12	ADQ88890
28	541.5	17.9	15185	4	AAK65589
29	525	17.4	329	6	ABL40824
30	511.5	16.9	15181	4	AAK65588
31	437.5	14.5	564	8	ABT23453
32	343.5	11.4	382	5	AAAD14428
33	331.5	11.0	2830	4	ABL02327
34	329.5	10.9	2609	9	ACF35860
35	310.5	10.3	2020	4	ABL03297
36	310.5	10.3	2629	9	ACF35859
37	310.5	10.3	4020	4	ABL03296
38	302	10.0	167	6	ABL40825
39	298.5	9.9	2422	5	AA585331
40	295.5	9.8	2682	12	ADQ83707
41	295.5	9.8	2682	13	ADQ87011
42	295.5	9.8	2875	5	ABX71285
43	288.5	9.5	1857	6	ABK90199
44	288.5	9.5	2380	5	AA514817
45	288.5	9.5	2380	6	ABL59533

ALIGNMENTS

RESULT 1

ADA05679	ADA05679 standard; cDNA; 1740 BP.
XX	AC ADA05679;
XX	DT 06-NOV-2003 (first entry)
XX	Human NOV9a encoding cDNA SEQ ID NO:39.
XX	human; NOV9a; antidiabetic; anorectic; antibacterial; virucide; immunomodulator; cytosolic; neurotropic; neuroprotective; antiparkinsonian; antilipemic; gene therapy; human disease; metabolic disorder; diabetes; obesity; infection; cachexia; cancer; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.
XX	OS Homo sapiens.
XX	Key Location/Qualifiers
FT CDS	76..1689
FT	/*tag= a
FT	/product= "NOV9a"
XX	WO2003029424-A2.
XX	10-APR-2003.
XX	02-OCT-2002; 2002WO-US031373.

XX PR 02-OCT-2001; 2001US-0326483P.
XX PR 05-OCT-2001; 2001US-0327435P.
XX PR 05-OCT-2001; 2001US-0327449P.
XX PR 09-OCT-2001; 2001US-0327917P.
XX PR 09-OCT-2001; 2001US-0328029P.
XX PR 09-OCT-2001; 2001US-0328044P.
XX PR 09-OCT-2001; 2001US-0328056P.
XX PR 12-OCT-2001; 2001US-0328849P.
XX PR 15-OCT-2001; 2001US-0329414P.
XX PR 17-OCT-2001; 2001US-0330142P.
XX PR 18-OCT-2001; 2001US-0330309P.
XX PR 22-OCT-2001; 2001US-0341058P.
XX PR 24-OCT-2001; 2001US-0339266P.
XX PR 24-OCT-2001; 2001US-0343629P.
XX PR 29-OCT-2001; 2001US-0349575P.
XX PR 01-NOV-2001; 2001US-0346337P.
XX PR 17-APR-2002; 2002US-0373260P.
XX PR 19-APR-2002; 2002US-0373815P.
XX PR 19-APR-2002; 2002US-0373817P.
XX PR 19-APR-2002; 2002US-0373826P.
XX PR 19-APR-2002; 2002US-0373884P.
XX PR 22-APR-2002; 2002US-0374977P.
XX PR 16-MAY-2002; 2002US-0381037P.
XX PR 16-MAY-2002; 2002US-0381038P.
XX PR 17-MAY-2002; 2002US-0381042P.
XX PR 28-MAY-2002; 2002US-0383656P.
XX PR 29-MAY-2002; 2002US-0383831P.
XX PR 25-JUN-2002; 2002US-0391335P.
XX PR 01-OCT-2002; 2002US-00262511.
XX PA (CURA-) CURAGEN CORP.
XX PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
XX PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
XX PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
XX PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CBA, Shenoy SG;
XX PI Shimkets RA, Rothenberg WF, Leach MD, Agee ML, Berghs C, Dippippo VA;
XX PI Eiben AJ, Gangolli EA, Rieger DK, Spaderna SK;
XX PR WPI; 2003-381626/36.
XX PR P-PSDB; ADA05680.
XX PR New NOVX polypeptides and nucleic acids, useful for diagnosing,
XX PR preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
XX PR cancer or dyslipidemia, and in chromosome mapping, tissue typing or
XX PR pharmacogenomics.
XX PS Claim 20; Page 134; 586pp; English.
XX CC The present invention describes NOVX proteins, where X can be 1 to 55
XX CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
XX CC described above and a carrier; (2) a kit comprising, in one or more
XX CC containers, the composition described above; (3) an isolated nucleic acid
XX CC molecule which encodes a NOVX protein of the invention; (4) a vector
XX CC comprising the nucleic acid molecule described above; (5) a cell
XX CC comprising the above vector; (6) an antibody that immunospecifically
XX CC binds to the polypeptide described above; (7) methods for determining the
XX CC presence or amount of the above polypeptide or nucleic acid molecule in a
XX CC sample; (8) methods for determining the presence of or predisposition to
XX CC a disease associated with altered levels of expression of the above
XX CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
XX CC method of identifying an agent that binds to the polypeptide described
XX CC above; (10) a method for identifying a potential therapeutic agent for
XX CC use in treating a pathology that is related to an aberrant expression or
XX CC aberrant physiological interactions of the polypeptide; (11) a method of
XX CC screening for a modulator of activity or of latency or predisposition to
XX CC a pathology associated with the polypeptide; (12) a method for modulating
XX CC the activity of the polypeptide described above; (13) methods of treating
XX CC or preventing a pathology associated with the above polypeptide in a
XX CC mammal; and (14) a method for producing the above polypeptide. NOVX
XX CC sequences have antidiabetic, anorectic, antibacterial, virucide,

CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian
CC and antilipaeic activities, and can be used in gene therapy. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease. The polypeptide or the nucleic
CC acid molecule may be used to diagnose, treat or prevent metabolic
CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence encodes a human NOVX protein from
CC the present invention.
XX SQ Sequence 1740 BP; 389 A; 476 C; 530 G; 345 T; 0 U; 0 Other;

Alignment Scores: 3.33e-235 Length: 1740
Pred. No.: 3025.00 Matches: 562
Score: 3025.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-631-958-11 (1-562) x ADA05679 (1-1740)

Qy 1 HisGluAlaAlaAenGlyProAlaProLeuGlyValArgAlaProProAlaTtrpArgThr 20
Db 1 CACGAGCCCGCTAACGGTCGCGGCCCTCTGGCGTCGCGGCCGCCAGCTGGCGGACG 60

Qy 21 SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrp 40
Db 61 AGCCCGCGCGGAGATGGGGCGACGCGGGCGCGGAGCGCGTCAATCCGTCTGTGG 120

Qy 41 ValysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuTrpTrp 60
Db 121 GTGAAGCAGCAGCGCTCGCGCGTGGAGCCCGCGGGCTCTGCTGCGCTGTGGTGG 180

Qy 61 ArgSerProGlyProGlyAlaGlyAlaProGlyValaAlaAspAlaCysSerValProValSer 80
Db 181 CGGAGCCCGGGCGCGGAGCGCGGCCCGCGTGTGCTGCTGCTGCTGCTGCTGCTATCT 240

Qy 81 GluIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100
Db 241 GAGATCATCGCGTTGAGGAAACAGACGTTTCACGGGAAACATCAAGGCAGTGGAAAAATGG 300

Qy 101 GlnLysMetGlnLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHis 120
Db 301 CAGAAAAATGAAAAAGCCTTACGCTTTTACAGTTTCTGTTAAAGAGACGACGCGGCAC 360

Qy 121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrp 140
Db 361 CGCTGGGAAGTGGCGCAGGTGACTTCTGCTGTCCAGAGGAGCAGCTGTGTCACTTGTGG 420

Qy 141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal 160
Db 421 CTGCAGACCCCTGCGGAGATGCTGGAGAAGCTCAGCTCCAGACCAAAAGCATTTACTGGTA 480

Qy 161 PheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAla 180
Db 481 TTTATCAACCCGTTTGAGGAAAGGACAAGCGGATATATGAAAGAAAAAGTGGCA 540

Qy 181 ProLeuPheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGln 200
Db 541 CCACCTGTTACCTTAGCCTCCATCACCACCTGACATCATCGTTACTGAACATGCTAATCAG 600

Qy 201 AlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 220
Db 601 GCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGCATCGTCTGTGTGGC 660

Qy 221 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla 240
Db 661 GGAGATGGTATGTTTCAGCGAGGTCTGCACGGTCTGATTGGGAGACGACGAGAGCGGCC 720

Qy	241	GlyValAspGlnAsnHisProhArgAlaValLeuValProSerSerLeuArgIleGlyIle	260
Db	721	GGGGTCGACAGAACACCCCGGGCTGTGGTCCCGAGTAGCTCCGATGGAAATC	780
Qy	261	IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu	280
Db	781	ATTCGCGCAGGGTCAACGAGCTCGTGTGTACTCCAGCTGGGACACAGCGACGACAA	840
Qy	281	ThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHis	300
Db	841	ACCTCGGGCTGCATATGCTTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCAC	900
Qy	301	HisAsnSerThrLeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAsp	320
Db	901	CACACAGCACACTCTTCGTACTCCGTGTCCCTGCTGGGCTACGGTCTTACGGGAC	960
Qy	321	IleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGly	340
Db	961	ATCATCAGGACAGTGAAGAACGGTGTGGTCTTGCAGATACGACTTTTCAGGT	1020
Qy	341	LeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGln	360
Db	1021	TTAAAGACCTTCCTCCACCACCTGCTATGAAGGCACAGTGTCTCTCCTCGTCACAA	1080
Qy	361	HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg	380
Db	1081	CACCGGTGGATCTCCAAAGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTGCAGG	1140
Qy	381	GlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAla	400
Db	1141	CNAAGCAGCAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGTTTGAAGCTCGG	1200
Qy	401	GluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsn	420
Db	1201	GAGGACGTGGAGAGTGGCAAGTCTCTGTGGAAAGTTTCTGGCCATCAATGCCACAAAC	1260
Qy	421	MetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp	440
Db	1261	ATGCTCTGTGCTTGTGCGCGGAGCCCGAGGGCCCTCTCCCGGCTGCCCACTTGGGAGAC	1320
Qy	441	GlySerSerAspLeuIleLeuLysCysSerArgPheAsnPheLeuArgPheLeu	460
Db	1321	GGGTCTTTGACCTCATCTCATCCGGAATGTCTCAGGTTCAATTTTCTGAGATTCTC	1380
Qy	461	IleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluValTyrArgVal	480
Db	1381	ATCAGGCACACCAACAGCAGGACAGTTTGACTTCACTTTTGTGTAAGTTTATCGCGTC	1440
Qy	481	LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly	500
Db	1441	AAGAAATTCAGTTTACGTTCGAAGCACATGGAGGATGAGGACAGCGACCTCAAGGAGGGG	1500
Qy	501	GlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSer	520
Db	1501	GGGAAGAAGCGTTTGGGCACATTTGCAGCAGCCACCCCTCTGCTGCTGACCGCTCTCC	1560
Qy	521	AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVal	540
Db	1561	AACAGCTCTTGAACTGCGAGGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCCAGATC	1620
Qy	541	HisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSer	560
Db	1621	CACCTGCCAGCTGGTTCGACTCTTTGCAGGAGGAATTGAAGAGAAATCCGAAGCCAGACTCA	1680
Qy	561	HisSer 562	
Db	1681	CACAGC 1686	
AC	ABL40828	standard; cDNA; 4413 BP.	

XX	03-JUL-2002	(first entry)	
DT	Human sphingosine kinase-like protein encoding cDNA.		
DE	Human sphingosine kinase-like protein; intracellular signalling; gene;		
KW	cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;		
KW	autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	CDS	1..1689	
FT		/*tag= a	
FT		/product= "sphingosine kinase-like protein"	
FT		/note= "see ABB07857"	
FT	CDS	76..1689	
FT		/*tag= b	
FT		/product= "sphingosine kinase-like protein"	
FT		/note= "see ABB07856"	
XX	WO200228906-A2.		
XX	11-APR-2002.		
XX	05-OCT-2001; 2001WO-EP011516.		
PF	06-OCT-2000; 2000US-0238005P.		
PR	23-AUG-2001; 2001US-0314113P.		
XX	(FARB) BAYER AG.		
XX	Kossida S, Encinas J;		
XX	WPI: 2002-340094/37.		
DR	P-PSDB; ABB07856, ABB07857.		
XX	New reagent for modulating the activity of sphingosine kinase-like		
PT	protein polypeptide or polynucleotide and treating cancer, asthma,		
PT	allergy, an autoimmune disease, or a central or peripheral nervous system		
PT	disorder.		
XX	Claim 1; Fig 9; 120pp; English.		
XX	The invention relates to a human sphingosine kinase-like protein. The		
CC	polypeptide can be expressed by standard recombinant methodology. The		
CC	sphingosine kinase-like protein and gene can be used to regulate		
CC	intracellular signalling and consequently cell proliferation and		
CC	apoptosis. Such regulation is useful for treating cancer, allergies (e.g.		
CC	asthma), autoimmune diseases (e.g rheumatoid arthritis) and central and		
CC	peripheral nervous system disorders (e.g. Parkinson's disease). The		
CC	present sequence represents the human sphingosine kinase-like protein		
CC	encoding cDNA		
XX	Sequence 4413 BP; 1020 A; 1107 C; 1207 G; 1079 T; 0 U; 0 Other;		
SQ			
Alignment Scores:			
Pred. No.:	1.14e-234	Length:	4413
Score:	3025.00	Matches:	562
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-631-958-11	(1-562) x ABL40828	(1-4413)	
Qy	1	HisGluAlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThr	20
Db	1	CACGAGGCGCGTAAACGGTCCGCGCCCTCGCGCTCCGCGCCCGCCAGCTGGCGGACG	60
Qy	21	SerProAlaAlaGluMetGlyAlaThrGlyValAlaGluProLeuGlnSerValLeuTrp	40
Db	61	AGCCCGCGCGGAGATGGGGCGGCGGCGCGCGCGCGCTGCAATCGCTGTGTGG	120

```
Qy 41 VallysGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTrp 60
Db 121 GTGAAGCAGCAGCGCTGCGCGCTGAGCTGGAGCCCGCGGGCTCTGCTGCGCTGGTGG 180
Qy 61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSer 80
Db 181 CGGAGCCCGGGCCCGAGCGCGCGCCCGCGCGCGATGCTGCTGCTGCTGCTGCTATCT 240
Qy 81 GluIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100
Db 241 GAGATCATCGCCGTGAGGAACAGACGTTTACCGGAAACATCAAGGACGTGGAAAAATGG 300
Qy 101 GlnLysMetGlnLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHis 120
Db 301 CAGAAATGGAAGACCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCGGCAC 360
Qy 121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGlnGlnLeuCysHisLeuTrp 140
Db 361 CGCTGGAAGTGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGG 420
Qy 141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal 160
Db 421 CTGCAGACCCCTGCGGAGATGCTGGGAAGCTGACGTCCAGACCAAGCATTTACTGGTA 480
Qy 161 PheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAla 180
Db 481 TTTATCAACCCGTTGGAGNAAGGACNAGCGAAGCGGATATATGAAGAAAGTGGCA 540
Qy 181 ProLeuPheThrLeuAlaSerIleThrThraspileleValThrGluHisAlaAsnGln 200
Db 541 CCACTGTTCACTTAGCCTCCATCACCACATGACATCGTTACTGAACATGCTTAATCAG 600
Qy 201 AlalysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 220
Db 601 GCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTGTGTGCGG 660
Qy 221 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla 240
Db 661 GGAGATGGTATGTTACGCGAGGTGCTGCACGGTCTGATTTGGAGGACGACGAGGCGCC 720
Qy 241 GlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 260
Db 721 GGGGTCCAGCAAGAACCAACCCCGGGCTGTGTGGTCCCGCAGTAGCTCCGATTTGGAATC 780
Qy 261 IleProIaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu 280
Db 781 ATTCGGCGGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCAC 840
Qy 281 ThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHis 300
Db 841 ACCTCGGCGTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCAC 900
Qy 301 HisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp 320
Db 901 CACAACAGCACACTCTCTCGTACTCCGCTGCTCCCTGTGGCTACGGCTTCTACGGGGAC 960
Qy 321 IleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGly 340
Db 961 ATCATCAAGGACAGTGAGAAGAAACGGTGGTGGTCTTTCAGATACGACTTTTCAGGT 1020
Qy 341 LeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGln 360
Db 1021 TTAAGACCTTCTCTCCACCATGTGTATGAAGGACAGTGTCTTCTCTCCCTGCACAA 1080
Qy 361 HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg 380
Db 1081 CACACGGTGGATCTCCAAAGGATAGGAAGCCCTGCCGGGAGGATGCTTTGTTTGCAGG 1140
Qy 381 GlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAla 400
Db 1141 CAAAGCAAGCAGCAGCTGGAGGAGGAGCAAGAAAGCACTGTATGTTTGGAAAGCTGCG 1200
```

```
Qy 401 GluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsn 420
Db 1201 GAGCAGTGGAGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAAC 1260
Qy 421 MetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp 440
Db 1261 ATGTCTGTGTTGTGCGCGAGCCCGAGGGGCTCTCCCGGCTGCCCATTTGGGAGAC 1320
Qy 441 GlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeu 460
Db 1321 GGGTCTTCTGACCTCATCTCATCCGGAATGTCTCAGGTTCAATTTTCTGAGATTTCTC 1380
Qy 461 IleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgVal 480
Db 1381 ATCAGGCACACCAACACAGCAGGACCATGTTGACTTCACTTTTTGTGAAGTTTATCGCGTC 1440
Qy 481 LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly 500
Db 1441 AAGAAATTCAGTTTACGTGCAAGACATGAGAGGATGAGGACAGCGACCTCAAGGAGGGG 1500
Qy 501 GlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSer 520
Db 1501 GGGAAAGAGCGCTTGGGCACATTTGCAGCAGCACCCCTCTCTGCTGCTCACGCTCC 1560
Qy 521 AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVal 540
Db 1561 AACAGCTCTCTGGAACCTGCGACGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCA 1620
Qy 541 HisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSer 560
Db 1621 CACTGCCAGCTGTGTTGCACTCTTTTGACGAGGAATTAAGAGAATCCGAAGCCAGACTCA 1680
Qy 561 HisSer 562
Db 1681 CACAGC 1686
RESULT 3
ADJ96598
ID ADJ96598 standard; DNA; 4429 BP.
XX
AC ADJ96598;
XX
XX 06-MAY-2004 (first entry)
XX
DE Human lipid kinase KIAA1646 DNA SeqID 55.
XX
KW gene; ds; kinase; human; SNP; single nucleotide polymorphism;
KW tyrosine protein kinase; serine/threonine protein kinase; PK; STK;
KW gene therapy; cancer; immune-related disease; cardiovascular disease;
KW brain; neuronal associated disease; metabolic; inflammatory disorder;
KW cytosolic; neuroprotective; immunomodulator; antiinflammatory;
KW lipid kinase; KIAA1646.
XX
OS Homo sapiens.
OS 39.
XX
PH Key Location/Qualifiers
FT variation /*tag= a replace(2391,g)
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= b replace(2577,g)
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= c replace(3020,c)
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= d replace(3769,c)
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= e replace(4272,g)
FT /*standard_name= "Single nucleotide polymorphism"
XX
```

PN WO2004006838-A2.
 XX 22-JAN-2004.
 XX 15-JUL-2003; 2003WO-US021730.
 XX 15-JUL-2002; 2002US-0395632P.
 XX (SUGEN-) SUGEN INC.
 XX Whyte D, Manning G, Caenepeel S;
 XX WPI: 2004-122753/12.
 DR P-PSDB; ADJ96664.
 XX
 PT New nucleic acid molecule encoding a kinase polypeptide, useful for
 PT preparing a composition for treating diseases or disorders, e.g., cancer,
 PT or neurological, immunological or inflammatory disorders.
 XX
 PS Example 1; SEQ ID NO 55; 366pp; English.
 XX
 CC This invention relates to a novel isolated, enriched or purified nucleic
 CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
 CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
 CC as well as protein kinase-like enzymes. The present invention describes
 CC screening methods to identify agonists, antagonists and antibodies that
 CC can be used to modulate the activity or function of the mammalian kinase
 CC enzymes. As such, these compositions can be used for gene therapy
 CC purposes to treat diseases or disorders including cancer, immune-related
 CC diseases, cardiovascular disease, brain or neuronal associated disease,
 CC metabolic and inflammatory disorders. Accordingly, they exhibit
 CC cytosolic, neuroprotective, immunomodulator and anti-inflammatory
 CC activities. This polynucleotide sequence is a human kinase DNA sequence
 CC of the invention.
 XX
 SQ Sequence 4429 BP; 1020 A; 1115 C; 1213 G; 1081 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 1.3e-233 Length: 4429
 Score: 3012.00 Matches: 560
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.5% Indels: 0
 DB: 12 Gaps: 0
 US-10-631-958-11 (1-562) x ADJ96598 (1-4429)
 QY 3 AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProAlaProAlaTrpArgThrSerPro 22
 DB 23 GCCGCTAACGGTCCGGCGCCCTTCGGCGTCCGGCGCCCGCCCGCCCGCCCGCCCGCCG 82
 QY 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpVallys 42
 DB 83 GCGCGGGAGATGGGGCGACGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGG 142
 QY 43 GlnGlnArgCysAlaValSerLeuGluProAlaAlaGlnLeuLeuArgTrpArgSer 62
 DB 143 CAGCAGCGCTGCGCGCTGAGCTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 202
 QY 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluile 82
 DB 203 CCGGGGCGGGAGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 262
 QY 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys 102
 DB 263 ATCGCGGTGAGAAACACAGCTTTCACGGGAAACATCAAGGCGGTGGAATGGCAGAA 322
 QY 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrp 122
 DB 323 ATGGAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGCGCGCGCGTGG 382
 QY 123 LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuLeuCysHisLeuTrpLeuGln 142

Db 383 AAGTGGCGCAGGTGACTTTCTGTGTCCAGAGAGCAGCTGTGTCTACTTGTGGCTGCAG 442
 QY ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheile 162
 Db 443 ACCCTGGCGGAGATGCTGGAGAAGCTGACGTCCAGACCAGCAATTTACTGGTATTATC 502
 QY AsnProPheGlyGlyLysGlyGlnGlyLysArgileTyrGluArgLysValAlaProLeu 182
 Db 503 AACCCGTTTGGAGGAAAGGCAAGGCAAGCGGATATATGAAGAAAAAGTGGCCACTG 562
 QY PheThrLeuAlaSerIleThrAspIleileValThrGluHisAlaAsnGlnAlaLys 202
 Db 563 TTCACCTTAGCTCCATCACCACCTGACATCATCTGTACTGAACATGCTAATCAGGCCAAG 622
 QY GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLys 222
 Db 623 GAGACTCTGTATGAGATTAACATAGACAAATACGACGCGCATCGTGTGTGGCGGAGAT 682
 QY GlyMetPheSerGluValLeuHisGlyLeuileGlyArgThrGlnArgSerAlaGlyVal 242
 Db 683 GGTATGTTACGCGAGGTGCTGCACGGTCTGATTGGAGGACGACAGAGAGCGCGGGTTC 742
 QY AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleilePro 262
 Db 743 GACCAAGAACACACCCCGGGCTGTGTGTCCAGTAGCTCCGGATTGGAAATCATTC 802
 QY AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282
 Db 803 GCAGGGTCAACGGACTGCGTGTGTGTACTCCACCGTGGGCACGACGACGACGACAACTCG 862
 QY AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn 302
 Db 863 CGCTGTCATATCGTTGTTGGGACCTCGCTGCCCATGGATGTCTCTAGTCCACCAAC 922
 QY SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleile 322
 Db 923 AGCACACTCTTCGCTACTCGTGTCTGTCTGTGGCTTACGGCTTCTACGGGACATCATC 982
 QY LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342
 Db 983 AAGGACAGTGAAGAAACCGTGTGTGGTCTTGCAGATACGACTTTTTCAGGTTTAAAG 1042
 QY ThrPheLeuSerHisHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThr 362
 Db 1043 ACCTTCTCTCTCCACCATGCTATGAAGGGACAGTGTCTCTCTCTCTCTCTCTCTCT 1102
 QY ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382
 Db 1103 GTGGATCTCCAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGAGGCGCAAGC 1162
 QY LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402
 Db 1163 AAGCAGCAGCTGGAGGAGGACAGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGAC 1222
 QY ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 422
 Db 1223 GTGGAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC 1282
 QY CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442
 Db 1283 TGTGCTTGTCCCGGAGGCCCGGGCGCTCTCCCGGCTGCCCATCTTGGGAGAGCGGGTCT 1342
 QY SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg 462
 Db 1343 TCTGACCTCATCTCATCCGGAATGTCTCAGGTTCAATTTTCTGAGATTTCTCATCAGG 1402
 QY HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys 482
 Db 1403 CACACCAACGACGAGGACAGTCTTCTACTTCTCTTGTGGAAGTTTATTCGGTCAAGAAA 1462
 QY PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys 502
 Db 1463 TTCCAGTTTACGTGGAAGCATATGAGGATGAGGACGACGACCTCAAGAGGGGGGGAAG 1522

Qy 503 LysArgPheGlyHisIleCysSerHisProSerCysCysThrValSerAsnSer 522
 Db 1523 AAGCGCTTGGGCACATTTCAGCAGCACCCTCTCTGCTGTCACCGTCTCCACAGC 1582
 Qy 523 SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys 542
 Db 1583 TCCTGGAACTCGACGGGGAGGTCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCACTGC 1642
 Qy 543 GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 562
 Db 1643 CAGCTGGTTCAGTCTCTTTCACAGAGGAATTGAAGAGAATTCGAAGCCAGACTCACAGC 1702

RESULT 4

ADP55247

ID ADP55247 standard; cDNA; 4445 BP.

XX AC ADP55247;

XX DT 18-NOV-2004 (first entry)

XX DE Human PRO cDNA sequence SEQ ID NO:1223.

XX KW human; PRO; immune related disease; inflammatory immune response;
 KW immune response stimulation; antiallergic; antianaemic; antiarthritic;
 KW antirheumatic; antidiabetic; antinflammatory; antipsoriatic;
 KW antirheumatic; antichyroid; CNS; dermatological; gastrointestinal;
 KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
 KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
 KW virucide; gene therapy; gene; ss.

XX OS Homo sapiens.

XX PN WO2004039956-A2.

XX PD 13-MAY-2004.

XX PF 28-OCT-2003; 2003WO-US034381.

XX PR 29-OCT-2002; 2002US-0422472P.

XX PA (GETH) GENENTECH INC.

XX PI Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;

XX PI Wood WL, Wu TD;

XX DR WPI; 2004-376182/35.

XX DR P-PSDB; ADP55248.

XX PT New PRO polynucleotides and polypeptides, useful in useful in diagnosing
 PT and treating an immune related disease, e.g. systemic lupus
 PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
 PT stimulating an immune response.

XX FS Claim 2; SEQ ID NO 1223; 3009pp; English.

XX CC The present invention describes an isolated PRO nucleic acid (1). Also
 CC described: (1) a vector comprising (1); (2) a host cell comprising the
 CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an
 CC isolated PRO polypeptide; (5) a chimeric molecule comprising the
 CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
 CC antibody which specifically binds to a polypeptide of (4); (7) a
 CC composition of matter comprising a polypeptide of (4), an agonist or
 CC antagonist of the polypeptide or an antibody that binds to the
 CC polypeptide in combination with a carrier; (8) an article of manufacture
 CC comprising a container, a label on the container and a composition of
 CC matter of (7); (9) a method of treating an immune related disease in a
 CC mammal; (10) a method for determining the presence of a PRO polypeptide
 CC in a sample suspected of having the polypeptide; (11) a method of
 CC diagnosing an immune related disease or an inflammatory immune response
 CC in mammal; (12) a method of identifying a compound that inhibits or
 CC mimics the activity of or expression of a gene encoding a PRO polypeptide
 CC ; and (13) a method of stimulating the immune response in a mammal. The

CC PRO sequences have antiallergic, antianaemic, antiarthritic,
 CC antirheumatic, antidiabetic, antinflammatory, antipsoriatic,
 CC antirheumatic, antichyroid, CNS, dermatological, gastrointestinal,
 CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
 CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
 CC virucide activities, and can be used in gene therapy. The nucleic acid
 CC (1) and the encoded polypeptides, compositions, kits and methods are
 CC useful in diagnosing and treating an immune related disease and in
 CC stimulating an immune response. The present sequence represents a human
 CC PRO nucleotide sequence from the present invention.

XX SQ Sequence 4445 BP; 1018 A; 1134 C; 1216 G; 1077 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-31e-233 Length: 4445
 Score: 3012.00 Matches: 560
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.57% Indels: 0
 DB: 13 Gaps: 0

US-10-631-958-11 (1-562) x ADP55247 (1-4445)

Qy 3 AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerPro 22

Db 55 GCGCGCTAACCGTCCGGCGCCCTCGCGCTCCGGCGGCCCGCCAGCTGGCGGAGCGAGCCG 114

Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys 42

Db 115 GCGGCGAGATGGGGGCGAGCGGGGGCGGAGCCGCTGCAATCCGTCGTGGGTGAAG 174

Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer 62

Db 175 CAGCAGCGCTGCCCGCTGAGCGCTGGAGCCCGCGCGGCGCTCTGCTGCGCTGGCGGAGC 234

Qy 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle 82

Db 235 CCGGGCGCGGAGCGCGCCCGCGCGGAGTGCCTGCTCTGCTGCTGTATCTGAGATC 294

Qy 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys 102

Db 295 ATGCGCGTTGAGGAACACAGACGTTACGCGGAACATCAAGCAGTGGAAATGGCAGAA 354

Qy 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTyr 122

Db 355 ATGGAAGAGCGCTTACGCTTTTACAGTTCACTGTGTAAGAGAGCAGCAGCGCCGCTGG 414

Qy 123 LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGln 142

Db 415 AAGTGGCGCGAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTCTCACTTGTGGCTGCAG 474

Qy 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162

Db 475 ACCCTCGGGAGATGCTGGAGAAGCTGACGCTCAGACCAAGCATTTTACTGGTATTATC 534

Qy 163 AsnProPheGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu 182

Db 535 AACCCGTTTGGAGGAAAGGACCAAGCAAGCGGATATATGAAAGAAAGTGGCACCCTG 594

Qy 183 PheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLys 202

Db 595 TTCACCTTAGCCCTCCATCACCACCTGATCATCATCTGTAACATGCTAATTCAGGCCAAG 654

Qy 203 GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp 222

Db 655 GAGACTCTGATGAGATTAACTAGACAATAACAGCGCATGCTGCTGTGCGCGGAGAT 714

Qy 223 GlyMetPheSerGluValLeuHisGlyIleGlyArgThrGlnArgSerAlaGlyVal 242

Db 715 GGTATGTTACGCGAGGTGCTGCGCGGTCTGATTGGGAGGACGAGAGCGCGCGGGTTC 774

Qy 243 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIlePro 262

Db 775 GACGAGAACCCCGGGCTGTGCTGGTGTGCTCCAGTAGCTCCGATTGGATCATTTCCC 834
Qy 263 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282
Db 835 GCAGGGTCAACGACTGCGTGTGTACTCCACCGTGGGCACCGAGCGCAGAAACCTCG 894
Qy 283 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisIle 302
Db 895 GCGCTGCATATCGTTGTGGGACTCGCTGGCCATGGATGTCTCTAGTCCACCACAAC 954
Qy 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322
Db 955 AGCACACTCTTCGCTACTCCGTGTCTCTGCTGGGCTACGGCTCTACGGGGACATCATC 1014
Qy 323 LysAspSerGluLysIleArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342
Db 1015 AAGCAGTGTGAGAAACGGTGTGGTCTTTCGACATACGACTTTTCAGGTTTAAAG 1074
Qy 343 ThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 362
Db 1075 ACCTTCCTCTCCACACTGCTATGAAGGAGCAGTGTCTCTCTCCCTGCACACACAG 1134
Qy 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382
Db 1135 GTGGGATCTCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTGTGTGAGGCAAGC 1194
Qy 383 LysGlnGlnLeuGluGluGlnLysIleAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402
Db 1195 AAGCAGCAGCTGGAGGAGAGCAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGGAC 1254
Qy 403 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 422
Db 1255 GTGGAGGAGTGGCAAGTCTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC 1314
Qy 423 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442
Db 1315 TGTGCTTGTGCGCGAGCCCGAGGGCTCTCCCGGGTGGCCACTTGGGAGCGGGTCT 1374
Qy 443 SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg 462
Db 1375 TCTGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG 1434
Qy 463 HisThrAsnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482
Db 1435 CACACCAACACGACGACGAGTTTCACTTCTTGTGGAAGTTTATCGCGTCAAGAAA 1494
Qy 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys 502
Db 1495 TTCCAGTTTACGTCGACGACATGAGGATGAGGACGAGCCTCAAGAGGGGGGGAAG 1554
Qy 503 LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSer 522
Db 1555 AAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCTGCTGTGTCACCGTCTCCACAGC 1614
Qy 523 SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys 542
Db 1615 TCCTGGAATCGGATGGGAGGTCCTGCACAGCCCTGCCATCGAGGTGAGAGTCCACTGC 1674
Qy 543 GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 562
Db 1675 CAGCTGGTTCGACTCTTTCGACGAGGATTTGAGAGATTCGAGCCAGATCCACAGC 1734

RESULT 5

ABA96945

ID ABA96945 standard; cDNA; 4463 BP.

XX

AC ABA96945;

XX

DT 20-MAY-2002 (first entry)

XX

DE Human ceramide kinase hCERK1-encoding cDNA.

XX

KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;

KW neurological disease; inflammation; human immunodeficiency virus;
KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;
KW cancer; neuroprotective; anti-inflammatory; anti-HIV; antidiabetic;
KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;
KW gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FH CDS 124..1737

FT /*tag= a

XX /product= "Human ceramide kinase hCERK1"

FN WO200196575-A1.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001WO-JP004889.

XX 14-JUN-2000; 2000JP-00178039.

XX (SANY) SANKYO CO LTD.

XX Sugiura M, Kono K, Kohama T;

XX WPI; 2002-179513/23.

XX P-PSDB; AAM49115.

Human ceramide kinase gene and the enzyme encoded by it for screening
substances as drugs for neurological, inflammatory and other disorders.
Claim 5; Page 46-53; 61pp; Japanese.

This sequence represents cDNA encoding a human ceramide kinase designated
hCERK1. The invention relates to hCERK1, nucleic acids encoding it,
expression vectors and host cells containing hCERK1 nucleic acids, the
recombinant production of hCERK1 and antibodies specific for hCERK1. The
invention also encompasses methods of isolating hCERK1 from samples, the
use of hCERK1 in drug screening, and the use of hCERK1 nucleic acid
sequences in gene therapy. hCERK1 mediates the ATP-dependent 1-
phosphorylation of ceramides and can be used to screen for therapeutic
and preventive agents for a wide range of disorders. Such disorders
include neurological disease, inflammation, human immunodeficiency virus
(HIV) infection, type 2 diabetes, obesity, sepsis, arteriosclerosis and
cancer

SQ Sequence 4463 BP; 1026 A; 1135 C; 1217 G; 1084 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 5.84e-233 Length: 4463
Score: 3004.00 Matches: 558
Percent Similarity: 99.82% Conservative: 1
Best Local Similarity: 99.64% Mismatches: 1
Query Match: 99.31% Indels: 0
DB: 6 Gaps: 0

US-10-631-958-11 (1-562) x ABA96945 (1-4463)

Qy 3 AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerPro 22
Db 55 GCCGCTAACGGTCCGGCGCCCTCGGGTCCGCGGCCCGCCAGCTGGCGACGAGCCCG 114
Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys 42
Db 115 GCGCGGAGATGGGGCGACGCGGGGCGCGAGCCGCTGCAATCCGTGCTGTGGGTGAAG 174
Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer 62
Db 175 CAGCAGCGCTCGCCGCTGAGCTGAGCCCGCGCGGGCTCTGTGCGCTGTGGCGGAGC 234
Qy 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle 82
Db 235 CCGGGGCGCGAGCGCGGCCCGCCCGCGCGGATGCCCTGTCTGTGCTGTATGTGAGATC 294

```
Qy 83 IleAlaValGluGluThrAspValHisGlyIysHisGlnGlySerGlyIysTrpGlnLys 102
Db 295 ATCCCGCTTGAGGAAACAGAGCTTCACGGGAAACATCAAGGCACGTGGAAATGCGAGAA 354
Qy 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaAraGArgHisArgTrp 122
Db 355 ATGGAAGAGCCCTTACGGCTTTTACAGTTCACTGTGTAAAGAGAGACGACGCGCACCGCTGG 414
Qy 123 LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGln 142
Db 415 AAGTGGCGCAGGTGACTTTCGTGTGTCAGAGAGCAGCTGTGTCACTTGTGTGCTGCAG 474
Qy 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162
Db 475 ACCCTGCGGAGATGCTGGAGAACCTGACGTCCAGACCAAGCAATTTACTGGTATTATC 534
Qy 163 AsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluAraGlyValAlaProLeu 182
Db 535 AACCCGTTTGGAGGAAAGGACAAAGCAAGCGGATATATGAAGAAAGTGGCACCACTG 594
Qy 183 PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys 202
Db 595 TTCACCTTAGGCCTCCATCACCACATCATCATCGTTACTGAACATGCTAATTCAGGCCAAG 654
Qy 203 GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp 222
Db 655 GAGACTCTGTATGAGATTAAATAGACAAATACACGGCATCGTCTGTGTCGGCGGAGAT 714
Qy 223 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnAraGlyVal 242
Db 715 GGTATGTTACGAGAGGTGCTGCACGGTCTGATTGGGAGACGACAGAGCGCGCGGGTCT 774
Qy 243 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro 262
Db 775 GACCAGAACCAACCCCGGGCTGTGCTGTGCTCCAGTAGCTCCCGATGGGAATCATTTCCC 834
Qy 263 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282
Db 835 GCAGGTCACAGGACTGGGTGTGTACTCCACCGTGGGCACGACGACGACGAGAACCTCG 894
Qy 283 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn 302
Db 895 GCGCTGCATATCGTTGTTGGGACATCGCTGGCCATGGATGTGCTCCTCAGTCCACCAAC 954
Qy 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322
Db 955 AGCACCTCCTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGAATCATC 1014
Qy 323 LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342
Db 1015 AAGGACAGTGAGAGAAACGGTGGTGGGTCTTCCAGATACGACTTTTCAGGTTTAAAG 1074
Qy 343 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 362
Db 1075 ACCTTCCTCTCCCACTATGATGAAGGACAGTGTCTTCTCCCTCCCTGCACACACACG 1134
Qy 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382
Db 1135 GTGGGATCTCAAGGATAGGAAGCCCTGCCGGGAGAGATGCTTTGTTGGAGGCAAGC 1194
Qy 383 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402
Db 1195 AAGCAGCAGCTGGAGGAGGACAGAGAAACCACTGATGTGTTTGAAGCTGCGGAGGAC 1254
Qy 403 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 422
Db 1255 GTGGAGGAGTGGCAAGTCTGTGGGAAAGTTTCTGGCCATCAATGCCACAAACATGTCC 1314
Qy 423 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442
Db 1315 TGTGCTTGTCCCGAGAGCCCGGGGCTCTCCCGGCTGCCCCCTTGGGAGACGGGTCT 1374
```

```
Qy 443 SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg 462
Db 1375 TCTGACCTCATCTCTATCCGAAATGCTCCAAGTTCAATTTTCTGAGATTTCTCATCAGG 1434
Qy 463 HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482
Db 1435 CACACCAACACGAGGACCACTTTCACCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 1494
Qy 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLys 502
Db 1495 TTCAGTTTACGTCGAGACACATGGAGGATGAGACACGACCTCAAGAGGGGGGAG 1554
Qy 503 LysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSer 522
Db 1555 AAGCGCTTTGGGCACATTTTGCAGCAGCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1614
Qy 523 SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys 542
Db 1615 TCTTGGAACTGCGATGGGAGGTCTTCGACAGCCCTGCCATCGAGGTCCAGTCCACTGC 1674
Qy 543 GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 562
Db 1675 CAGCTGGTTCGACTCTTTGCACGAGGAAATTGAAGAGAAATCCGAAGCCAGACTCACAGC 1734
RESULT 6
ADN62844
ID ADN62844 standard; DNA; 1740 BP.
XX
AC ADN62844;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human NOV9a DNA.
XX
KW ds; gene; human; NOVX; metabolic disorder; diabetes; obesity;
KW infectious disease; anorexia; cancer; cancer-associated cachexia;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia;
KW metabolic syndrome X; wasting disorder.
XX
OS Homo sapiens.
XX
PN US2004038223-A1.
XX
PD 26-FEB-2004.
XX
PF 01-OCT-2002; 2002US-00262511.
XX
PR 02-OCT-2001; 2001US-0326483P.
PR 05-OCT-2001; 2001US-0327435P.
PR 05-OCT-2001; 2001US-0327449P.
PR 09-OCT-2001; 2001US-0327917P.
PR 09-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328044P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328849P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 18-OCT-2001; 2001US-0330309P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0339266P.
PR 24-OCT-2001; 2001US-0343629P.
PR 29-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 17-APR-2002; 2002US-0373260P.
PR 19-APR-2002; 2002US-0373815P.
PR 19-APR-2002; 2002US-0373817P.
PR 19-APR-2002; 2002US-0373826P.
PR 19-APR-2002; 2002US-0373884P.
PR 22-APR-2002; 2002US-0374977P.
PR 16-MAY-2002; 2002US-0381037P.
PR 16-MAY-2002; 2002US-0381038P.
PR 16-MAY-2002; 2002US-0381042P.
```


Db 720 CGGGTTCGACACCAACCCCGGCTGTGCTGCCAGTAGCCTCCGGATTGGAAT 779
Qy 260 eileProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaG1 280
Db 780 CATTCGCGCAGGGTCAACGGAATCGGTGTGTACTCCACCGTGGGACACGACGCGAGA 839
Qy 280 uThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHi 300
Db 840 AACCTCGGCGTCATATCGTTGTTGGGACTCGCTGGCCATGATGTGTCTCAGTCCA 899
Qy 300 sHisAsnSerThrLeuLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAs 320
Db 900 CCACACAGCACACTCTCTTCGCTACTCCGTGTCCGTGCTGGGCTTCTACGGGGA 959
Qy 320 pilelleIleAspSerGluLysLysLysValArgTyrLeuGlyLeuAlaArgTyrAspPheSerG1 340
Db 960 CATCATCAGGACAGTGAAGAAGCGGTGTGGTCTTCCAGATACGACTTTTCAGG 1019
Qy 340 YLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaG1 360
Db 1020 TTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACA 1079
Qy 360 nHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysAr 380
Db 1080 ACACACGGTGGGATCTCCAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGCAG 1139
Qy 380 gGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaA1 400
Db 1140 GCAAAGCAGCAGCAGCTGGAGGAGGACGAAGAAGCATGTATGTTTGGAGCTGC 1199
Qy 400 aGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAs 420
Db 1200 GGAGGACGTGGAGGATGGCAAGTCGTCTGGGAAGTTCTTGCGCCATCAATGCCACAA 1259
Qy 420 nMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAs 440
Db 1260 CATGTCCTGTCTGTGTCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCACTTGGGAGA 1319
Qy 440 pGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLe 460
Db 1320 CGGGTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTCAATTTCTCGAGATTCT 1379
Qy 460 uileArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgVa 480
Db 1380 CATCAGGCACACCAACACGACGAGCCAGTCTTGTGACTTCTTGTGAAGTTTATCGGT 1439
Qy 480 lLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluG1 500
Db 1440 CAAGAAATTCAGTTTACGTGGAAGCACATGGAGGATGAGGACGCGACCTCAAGGAGGG 1499
Qy 500 YGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSe 520
Db 1500 GGGGAAGAAGCGCTTTGGGCACATTTGACGAGCACCCCTCTGCTGCTGCACCGTCTC 1559
Qy 520 rAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVa 540
Db 1560 CAACAGCTCTCGGAATCGACGCGGGAGTCTTCGACAGCCCTGCCATCGAGGTCAGAGT 1619
Qy 540 lHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSe 560
Db 1620 CCATCTGCCAGCTGGTTGCACTCTTTGACGAGGAATTGAAGAGATCCGAAGCCAGACTC 1679
Qy 560 rHisSer 562
Db 1680 ACACAGC 1686

RESULT 7
ABX70921
ID ABX70921 standard; cDNA; 4432 BP.
XX
AC ABX70921;
XX
DT 05-MAR-2003 (first entry)

XX Novel human cDNA sequence #146.
DE Human; gene; ss; nervous system disorder; peripheral neuropathy;
KW Huntington's disease; amyotrophic lateral sclerosis; haemophilia;
KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
KW insulin-dependent diabetes mellitus; anaemia; thrombocytopenia; wound;
KW ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;
KW fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;
KW coagulation disorder; cancer; tumour; inflammatory disease; septic shock;
KW Crohn's disease; anaphylaxis; proliferation; chemotactic;
KW differentiation; stem cell growth factor; haematopoiesis; chemokinetic;
KW haemostatic; antiinflammatory; expressed sequence tag; EST.
XX Homo sapiens.
OS
XX
XX WO200281731-A2.
XX 17-OCT-2002.
XX
XX 29-JAN-2002; 2002WO-US001222.
XX
XX 30-JAN-2001; 2001US-00774528.
PR (HYSE-) HYSEQ INC.
PA (GOOD/) GOODRICH R W.
PA
XX
XX Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Zhouman T, Wang J, Wang D, Drmanac RT;
XX
XX WPI; 2003-058563/05.
XX
XX Novel polypeptide useful for treating neurodegenerative diseases, myeloid
PT or lymphoid cell disorders, bone disorders, mechanical and traumatic
PT disorders, coagulation disorders, and inflammatory diseases.
XX
XX Claim 1; Page; 612pp; English.
XX
XX This invention relates to the cDNA sequences encoding an isolated novel
CC human polypeptide. The protein encoded by the nucleic acid of the
CC invention is useful for treating central and peripheral nervous system
CC diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic
CC lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,
CC Alzheimer's disease); autoimmune disease (e.g. systemic lupus
CC erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)
CC ; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopenia)
CC ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,
CC osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head
CC trauma); lung or liver fibrosis; reperfusion injury in various tissues;
CC bacterial, viral or fungal infections; allergic conditions such as
CC allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);
CC cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's
CC disease, anaphylaxis). The protein may be used to inhibit the growth,
CC infection or function of infectious agents such as bacteria, fungi,
CC viruses, or to effect bodily characteristics, biorhythms or circadian
CC cycles of rhythms. The protein may also have
CC proliferation/differentiation, stem cell growth factor, haematopoiesis
CC regulation, immune stimulating or suppressing, chemotactic/chemokinetic,
CC haemostatic and thrombolytic, receptor/ligand, and antiinflammatory
CC activities. The cDNA sequences of the invention are useful for expressing
CC recombinant protein for analysis. The present sequence represents a novel
CC human cDNA sequence of the invention, this sequence is an expressed
CC sequence tag (EST) and was identified using subtractive hybridisation
XX
SQ Sequence 4432 BP; 970 A; 1122 C; 1282 G; 1058 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.33e-227 Length: 4432
Score: 2931.50 Matches: 548
Percent Similarity: 98.21% Conservative: 2
Best Local Similarity: 97.86% Mismatches: 7
Query Match: 96.91% Indels: 3

DB:	8	Gaps:	1
US-10-631-958-11 (1-562) x ABX70921 (1-4432)			
Qy	6	GlyProAlaProLeuGly-----ValArgAlaProProAlaTTPArgThrSerPro	22
Db	1135	GGTTCCCGCCCTATAGACAAGACAGTCAACAGGAGCTGGCAGGCTTGGCGACAGAGCCG	1194
Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTTPValLys	42
Db	1195	CGCGCGAGATGGGGCCGACGGGGCGGAGCGCTGCATCCGTCTGTGGGTGAAG	1254
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTTPArgSer	62
Db	1255	CAGCAGCGCTCGCGCGTGGAGCTGGAGCCGCGCGGCTCTCTGTGCGCTGTGTGCGGAGC	1314
Qy	63	ProGlyProGlyAlaGlyAlaProGlyValAspAlaCysSerValProValSerGluIle	82
Db	1315	CCGGGGCCGAGCGCGGCCCCCGGCGGATGCCCTCTGTGCTCTGTATCTGAATC	1374
Qy	83	IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTTPGlnLys	102
Db	1375	ATCCCGCTTGAGGAAACAGAGCTTCCGGGAAACATCAAGCAGTGGAAANTGGCAGAA	1434
Qy	103	MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTTP	122
Db	1435	ATGGAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGACCGCTGG	1494
Qy	123	LysTTPAlaGlnValThrPheTTPCysProGluGlnLeuCysHisLeuTTPLeuGln	142
Db	1495	AAGTGGCGCAGGTGACTTCTGTGTGTCAGAGAGCAGCTGTGTCACTTGTGGCTGCAG	1554
Qy	143	ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle	162
Db	1555	ACCTCGGGGAGATGCTGGAGAGCTGACGTCCAGACAAAGCATTTACTGTGTTATTC	1614
Qy	163	AsnProPheGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu	182
Db	1615	AAACCGTTTGGAGGAAAGGACAGGCAAGCGGATATATGAAGAAAGTGGCACCACTG	1674
Qy	183	PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys	202
Db	1675	TTCACCTTAGCCTCCATCACACATGACATCATCGTTACTGAACATGTCTAATACGGCCAG	1734
Qy	203	GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp	222
Db	1735	GAGACTCTGTATGAGATTAAATAGACAAATACGACCGCATCTGTCTGTGCGCGGAGAT	1794
Qy	223	GlyMetPheSerGluValLeuHisGlyIleGlyArgThrGlnArgSerAlaGlyVal	242
Db	1795	GGTATGTTACGAGAGTGTCTGACCGTCTGATTGGGAGGACGACAGGAGCGCGGGGTC	1854
Qy	243	AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro	262
Db	1855	GACCAGAACCAACCCCGGGTGTCTGTGTCCCAAGTAGCTCCGGATTGGAAATCATTTCC	1914
Qy	263	AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer	282
Db	1915	GCAGGGTCAACGACTCGTGTGTATCTCACCGGTGGGCAACGACGACGAGAAACCTCG	1974
Qy	283	AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn	302
Db	1975	CGCTGCATATCTGTTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCACAAC	2034
Qy	303	SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle	322
Db	2035	AGCACACTCTTCCTTACCTCCGTCTCCCTGTGGGCTACGCGCTTCTACGGGGACATCATC	2094
Qy	323	LysAspSerGluLysLysArgTTPLeuGluValAlaArgTyrAspPheSerGlyLeuLys	342
Db	2095	AAGGACAGTGAAGAAGACGGTGTGTGGGTTCTGCCAGATACGACTTTTTCAGGTTTAAAG	2154
Qy	343	ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr	362

Db	2155	ACCTTCTCTCCCAACCACCTGCTATGAGGAGCAGTGTCTTCTCTCCCTGTCACAAACACAGC	2214
Qy	363	ValGlySerProArgAspArgLysProCysArqAlaGlyCysPheValCysArgGlnSer	382
Db	2215	GTGGGATCTCCAAGGATAGGAAGCCCTCGCGGCGAGGATGCTTTGTTTGCAGCGAAAGC	2274
Qy	383	LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaGluAsp	402
Db	2275	AAGCAGCAGCTGAGGAGGAGCAGAGAAGCACTGTATGGTTTGGAAAGCTCGGAGGAC	2334
Qy	403	ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer	422
Db	2335	GTGGAGAGTGGCAAGTCTGCTCTGGGAAGTTTCTGCCCATCAATGCCCAAAACATGTCC	2394
Qy	423	CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer	442
Db	2395	TGTGCTTGTCCGCGAGAGCCCGAGGGCCCTCTCCCGGCTGCCCATCTGGGAGACGGGTCT	2454
Qy	443	SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg	462
Db	2455	TCTGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG	2514
Qy	463	HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys	482
Db	2515	CACACCAACAGCAGGACCAGATTGACTTTCACCTTTTGTGTAAGTTTATCGCGTCAAGAAA	2574
Qy	483	PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLys	502
Db	2575	TTCAGTTTACGTCGGAAGACATGAGGAGATGAGGACAGCGACTCAAGGAGGGGGGAAG	2634
Qy	503	LysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSer	522
Db	2635	AAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCTCTGCTGCGACCGTCTCCAACAGC	2694
Qy	523	SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys	542
Db	2695	TCCTGGAACTGCGACGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGC	2754
Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer	562
Db	2755	CAGCTGGTTCGACTCTTTGCACGAGGAATTGAAGAGAAATCCGGAAGCCAGACTCACACAGC	2814
RESULT 8			
AAD14426			
ID	AAD14426 standard; cDNA; 1840 BP.		
XX			
AC	AAD14426;		
XX			
XX	01-NOV-2001 (first entry)		
XX			
DE	Human sphingosine kinase (SphK) cDNA #2.		
XX			
KW	Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy;		
KW	antitense therapy; cancer; sphingolipid; signalling molecule; apoptosis;		
KW	cytostatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;		
KW	leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	109..1524		
FT	/*tag= a		
FT	/product= "Human sphingosine kinase (SphK) protein #2"		
XX			
PN	WO200160990-A2.		
XX			
PD	23-AUG-2001.		
XX			
XX	14-FEB-2001; 2001WO-US004789.		
XX			
PR	14-FEB-2000; 2000US-0182360P.		
PR	22-MAR-2000; 2000US-0191261P.		

XX (CURA-) CURAGEN CORP.
PA (GETH) GENENTECH INC.
XX Rastelli L;
XX WPI; 2001-514770/56.
DR P-PSDB; AAE07884.
XX An isolated Sphingosine kinase polypeptide useful for treating a Sphk-
PT associated disorder especially cancer, restenosis or ischemia in a human.
XX Claim 8; Page 94-95; 107pp; English.
XX The present invention relates to sphingosine kinase (SphK) polypeptides
CC and nucleic acids encoding them. SphK is useful for treating a Sphk-
CC associated disorder especially cancers such as leukaemia, lymphoma,
CC ovarian, breast, lung, colon, testicular, stomach and skin,
CC atherosclerosis, restenosis or ischaemia and cell proliferative disease
CC or disorder associated with vascular diseases. SphK gene is used in gene
CC therapy and antisense-therapy. Sphingolipids serving as signalling
CC molecules have recently emerged as regulators of cell growth,
CC differentiation, diverse cell phenotypes and cell death. Activation of
CC SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
CC endothelial cells. The present sequence is human sphingosine kinase
CC (SphK) cDNA
XX
SQ Sequence 1840 BP; 459 A; 468 C; 503 G; 410 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.51e-205 Length: 1840
Score: 2649.50 Matches: 495
Percent Similarity: 97.93% Conservative: 1
Best Local Similarity: 97.63% Mismatches: 0
Query Match: 87.59% Indels: 11
DB: 5 Gaps: 1

US-10-631-958-11 (1-562) x AAD14426 (1-1840)

QY 67 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaValGlu 86
DB 1 GCCGGCCCCCGCGCGGATGCTCTCTGTGCTGTATCTGATCATCGCGTTGAG 60

QY 87 GluThrAspValHisGlyHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro 106
DB 61 GAAACAGAGCTTACGGGAAACATCAAGCGCAGTGGAAATGCGAATAATGGAAGCCT 120

QY 107 TyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAlaGln 126
DB 121 TACGCTTTTACAGTTCACTGTGTAAAGAGAGACGACGCGCACCGCTGGAAAGTGGCGCAG 180

QY 127 ValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu 146
DB 181 GTGACTTTTCTGGTGTCCAGAGAGAGCAGCTGTGTCTGTGGTGGCAGACCTCGCGGAG 240

QY 147 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 166
DB 241 ATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGGTATTATCAACCCGTTTGA 300

QY 167 GlyLysGlyGlnGlyLysArgIleTrpGluArgLysValAlaProLeuPheThrIleuAla 186
DB 301 GGAAGGACCAAGGCAAGCGGATATGAAAGAAAAGTGGCACCACCTGTTCACTTAGCC 360

QY 187 SerIleThrAspIleIle-----ValThr 195
DB 361 TCCATCACCCTGACATCATCGTTAAACAAATTTATGTATGTTAGAGTAAATTAAT 420

QY 196 GluHisAlaAsnGlnAlaLysGluThrLeuTrpGluIleAsnIleAspLysTrpAspGly 215
DB 421 GAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGC 480

QY 216 IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyIleGluGluArg 235
DB |||||

DB 481 ATCTCTGTGTGCGCGGAGATGGTATGTTTTCAGCGAGGTGTGTCACGGTCTGATTGGGAGG 540
QY 236 ThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSer 255
DB |||||
DB 541 ACCGAGAGAGCGCGCGGTCGACACAGAACACCCCGGGCTGTGCTGGTCCCCAGTAGC 600
QY 256 LeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTrpSerThrValGly 275
DB 601 CTCGGATTGGAATCATTTCCCGAGGGTCAACGAGCTGCGTGTGTATTCTCCACCGTGGC 660
QY 276 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAsp 295
DB 661 ACCAGACGACGAGAAACCTCGCGCTGCATGCTTGTGGGGACTCGCTGGCCCATGGAT 720
QY 296 ValSerSerValHisAsnSerThrLeuLeuArgTrpSerValSerLeuLeuGlyTyr 315
DB 721 GTGTCTCAGTCCACCAACACAGCACACTCTTTCGCTACTCCGTGTCTCTGGCTAC 780
QY 316 GlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArg 335
DB 781 GGCTTCTACGGGACATCATCAGGACAGTGAGAAGAAACGGTGGTGGGTCTTGGCCAGA 840
QY 336 TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTrpGluGlyThrValSer 355
DB 841 TACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAAGGACAGTGTCC 900
QY 356 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 375
DB 901 TTCTCTCTGCACACACACGCTGGGATCTCCAAGGATAGGAAGCCCTGCCGGCAGGA 960
QY 376 CysPheValCysArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyr 395
DB 961 TGCTTTTGTGACGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGCAACTGTAT 1020
QY 396 GlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAla 415
DB 1021 GGTTTGGAAGCTCGGAGGAGCGTGGAGAGTGGCAAGTGTCTGTGGAAAGTTCTGGCC 1080
QY 416 IleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAla 435
DB 1081 ATCAATGCCACAAACATGCTCTGTGCTTGTCCCGGAGCCCGAGGGCCTCTCCCCGCT 1140
QY 436 AlaHisLeuGlyAspGlySerSerAspLeuIleLeuLysLysCysSerArgPheAsn 455
DB 1141 GCCCAGCTGGGAGACGGGCTCTTCTGACCTCATCTCATCGGAAATGCTCCAGGTTCAAT 1200
QY 456 PheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheVal 475
DB 1201 TTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCAAGTTTGACTTTCACTTTGT 1260
QY 476 GluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSer 495
DB 1261 GAAGTTTATCGCTCAAGAAATTCAGTTTACGTGGAAGCACATGGAGGATGAGGACAGC 1320
QY 496 AspLeuLysGluGlyLysLysLysArgPheGlyHisIleCysSerSerHisProSerCys 515
DB 1321 GACCTCAAGGAGGGGGGAGAAAGCGCTTTGGGCACATTTTGCAGCAGCACCCCTCTGTC 1380
QY 516 CysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAla 535
DB 1381 TGCTGCAACCGTCTCCAAACAGCTCTGGAACCTGCGACCGGGAGGTCCTGACAGCCCTGCC 1440
QY 536 IleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsn 555
DB 1441 ATCGAGGTGAGGTTCCTGCTGCCAGCTGCTTTCGACTCTTTCACCGAGGAATGAAGAGAT 1500
QY 556 ProLysProAspSerHisSer 562
DB 1501 CCGAAGCGAGACTCACACAGC 1521

RESULT 9
AAAS0510
ID AAAS0510 standard; cDNA; 4231 BP.

XX AAA50510;
 AC
 XX
 DT 05-DEC-2000 (first entry)
 XX
 XX Human sphingosine kinase C cDNA.
 XX
 XX Sphingosine kinase C; SKC; human; drug screening; infection;
 KW antiinflammatory; antiallergic; anticancer; inflammation; allergy;
 KW cancer; therapy; diagnosis; ds.
 XX
 XX Homo sapiens.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 71..1453
 FT /*tag= a
 XX
 XX WO200052173-A2.
 PN
 XX
 XX 08-SEP-2000.
 PD
 XX
 XX 02-MAR-2000; 2000WO-CA000223.
 PF
 XX
 XX 02-MAR-1999; 99US-0122516P.
 PR
 XX
 XX (ALLX) NPS ALLELIX CORP.
 PA
 XX
 XX Munroe D, Gupta A, Falzone GR;
 PI
 XX
 XX WPI; 2000-572185/53.
 DR
 DR P-PSDB; AAY96059.
 XX
 XX New human sphingosine kinase A, B and C polynucleotides and polypeptides
 PT useful in e.g. chromosome and gene mapping, and detecting inflammation or
 PT disease associated with abnormal levels of sphingosine kinase expression.
 XX
 XX Disclosure; Fig 7; 81pp; English.
 PS
 XX
 CC The present sequence is that of an isolated polynucleotide encoding human
 CC sphingosine kinase C (SKC, see AAY96059), an enzyme that phosphorylates
 CC sphingosine to form sphingosine 1-phosphate. The polynucleotide was
 CC isolated from an HeLa cDNA library by PCR amplification. The invention
 CC provides polynucleotides (see AAA50508-10) and polypeptides (see AAY96057
 CC -59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC.
 CC The polynucleotides may be used as hybridization probes, in the
 CC construction of PCR primers for chromosome and gene mapping, in the
 CC recombinant production of SKA, SKB and SKC, and in the generation of
 CC antisense DNA or RNA. They can be used to detect inflammation or disease
 CC associated with abnormal levels of SK expression, or to detect
 CC differences in gene sequence between normal and carrier or affected
 CC individuals. Host cells expressing SK can be used in drug screening.
 CC Human SK specific antibodies, inhibitors, ligands or their analogues are
 CC useful as bioactive agents to treat inflammation or disease including
 CC viral, bacterial or fungal infections, allergic responses, mechanical
 CC injury associated with trauma, hereditary diseases, lymphoma or
 CC carcinoma, and other conditions with activate the genes of kidney, lung,
 CC heart, lymphoid or tissues of the nervous system
 XX
 SQ Sequence 4231 BP; 1022 A; 1021 C; 1086 G; 1067 T; 0 U; 35 Other;
 Alignment Scores:
 Ind. No.: 3,326-199 Length: 4231
 Score: 2587.00 Matches: 481
 Percent Similarity: 99.79% Conservative: 1
 Best Local Similarity: 99.59% Mismatches: 1
 Query Match: 85.52% Indels: 0
 DB: 3 Gaps: 0
 US-10-631-958-11 (1-562) x AAA50510 (1-4231)
 Qy 80 SerGluIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLys 99
 Db 2 TCTGATCATCGCCGTTGAGGAAACAGACGTTTACGGGAAACATCAAGGCAGTGGAAAA 61

Qy 100 TrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArg 119
 Db 62 TGGCAGAAAAATGGAAAAAGCCTTACGCTTTTACAGTTTACAGTTGTGTAAAGAGACGACGG 121
 Qy 120 HisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuLysHisLeu 139
 Db 122 CACCGCTGGAAAGTGGGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTG 181
 Qy 140 TrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeu 159
 Db 182 TGGCTGCAGACCCCTGCGGAGATGCTCGAGAAGCTGACGTCACAGACCAAGCAATTTACTG 241
 Qy 160 ValPheIleAsnProPheGlyGlyLysGlnClyLysArgIleLysGluArgLysVal 179
 Db 242 GTATTATTCAACCCGTTTGGAGGAAAAAGGCAAGCGCATATATGAAAGAAAAAGTG 301
 Qy 180 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsn 199
 Db 302 GCACCACTGTTACCTTAGCTCCATCCACNCTGACATCATCGTTACTGGAACATGCTAAT 361
 Qy 200 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 219
 Db 362 CAGGCCAAGGAGACTCTGTATGAGATTAAACATACAAATACGACGCGCATCGTCTGTGTC 421
 Qy 220 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSer 239
 Db 422 GCGGAGATGGTATGTTTCAGCGAGGTGCTGCACGCTGTGATTGGGAGACGCAGAGGAGC 481
 Qy 240 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 259
 Db 482 GCCGGGTGCACCAAGAACCCACCCCGGGCTGTGCTGTCCTCCAGTAGCTCCGATTGGA 541
 Qy 260 IleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAla 279
 Db 542 ATCATTTCCCGCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGCACGACGCGCA 601
 Qy 280 GluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerVal 299
 Db 602 GAAACCTCGCGCTGCATATCGTTTGGGAGCTCGCTGGCCATGGATGTCTCTCAGTC 661
 Qy 300 HisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGly 319
 Db 662 CACCACAAACAGCACACTCTTCGCTACTCGCTGTCCTGCTGGGCTACGGCTTCTACGGG 721
 Qy 320 AspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSer 339
 Db 722 GACATCATCAAGGACAGTGGAGAGAAACGGTGGTGGTCTTGGCCAGATACGACTTTTCA 781
 Qy 340 GlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAla 359
 Db 782 GGTTTAAAGACCTTCTCTCCACACTGCTATGAAGGGACAGTGTCTCTCTCCCTCCCTGCA 841
 Qy 360 GlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCys 379
 Db 842 CAACACACGGTGGATCTCCAAGGGATAGGAAGCCCTGCCGGCAGGAGATGTTGTTTGC 901
 Qy 380 ArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAla 399
 Db 902 AGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAGAACGACCTGTATGTTTGGAGCT 961
 Qy 400 AlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThr 419
 Db 962 GCGGAGGACGCTGGAGGAGTGGCAAGTCTCTGTGGGAAGTTTCTGGGCATCAATGCGCACA 1021
 Qy 420 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 439
 Db 1022 AACATGTCCTGTGCTGTGCGCGGAGCCCGCAGGGGCTCTCCCGGCTGCCCATTTGGGA 1081
 Qy 440 AspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPhe 459
 Db 1082 GACGGGTCTTCTGACCTCATCTCTCATCGGNAATGCTCCAGGTTCAATTTTCTGAGATT 1141

Qy 460 LeuLeuArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArg 479
Db 1142 CTCATCAGGCACACCAACAGCAGCAGCAGTTTGACTTCACTTTTGTGAAGTTTATCGC 1201
Qy 480 VallysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGlu 499
Db 1202 GTCAAGAAATTCAGTTTACGTCTGAAGCACATGATGAGGATGAGGACAGCAGCTCAAGGAG 1261
Qy 500 GLYGLYLYSLSVSAAGPheGlyHisIleCysSerSerHisProSerCysCysCysThrVal 519
Db 1262 GGGGGGAAGAGCCCTTTGGGCARTTTGCAGCAGCCCTCTCTGCTGCACCGTC 1321
Qy 520 SerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArg 539
Db 1322 TCCAAACAGCTCTGGAACTGCGACGGGAGGTCTCTGCACAGCCCTGCCATCGAGTCA 1381
Qy 540 ValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAsp 559
Db 1382 GTCCACTGCCAGCTGGTTTCGACTCTTTGCACGAGAAATTGGAAGAGAATCCGAAGCCAGAC 1441
Qy 560 SerHisSer 562
Db 1442 TCACACAGC 1450
RESULT 10
ID ADS10370 standard; DNA; 4702 BP.
XX ADS10370;
XX 16-DEC-2004 (first entry)
XX Human therapeutic DNA - SEQ ID 607.
XX antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
XX Homo sapiens.
XX WO2004080148-A2.
XX 23-SEP-2004.
XX 30-SEP-2003; 2003WO-US030720.
XX 02-OCT-2002; 2002US-0416186P.
XX (NUVE-) NUVELO INC.
XX Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX WPI; 2004-668857/65.
XX P-PSDB; ADS11054.
XX New polynucleotide, useful in preparing a composition for diagnosing or
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT aplastic anemia or cancer for promoting wound healing.
XX
XX Claim 1; SEQ ID NO 607; 718pp; English.
XX The invention relates to a novel isolated polynucleotide and the encoded
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during gene therapy
CC procedures. The current sequence is that of a human therapeutic DNA of
CC the invention. The current sequence is not shown explicitly within the
CC specification but can be accessed from the WIPO web-site.
XX

SQ Sequence 4702 BP; 970 A; 1226 C; 1364 G; 1118 T; 0 U; 24 Other;
Alignment Scores:
Pred. No.: 1,86e-195 Length: 4702
Score: 2541.50 Matches: 505
Percent Similarity: 65.72% Conservative: 3
Best Local Similarity: 65.33% Mismatches: 7
Query Match: 84.02% Indels: 259
DB: 13 Gaps: 4
US-10-631-958-11 (1-562) x ADS10370 (1-4702)
Qy 6 GlyProAlaProLeuGly-----ValArgAlaProProAlaTrpArgThrSerPro 22
Db 749 GGTTCGCGCCCTATAGACAGACAGTCAACAAGGAGCTGGCAGGCTTGGCGGAGAGCCCG 808
Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys 42
Db 809 GCGGCGAGATGGGGCGACGGGGCGGCGGAGCCGCTGCAATCCGCTGCTGTGGTGAAG 868
Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer 62
Db 869 CAGCAGCGCTGCCCGCTGAGCCTGGAGCCCGCGGGCTCTGTGCGCTGGTGGCGAGC 928
Qy 63 ProGlyProGlyAlaGlyAlaProGlyAla----- 72
Db 929 CCGGGGCGGAGCGCGCGCCCGCGGCGC-GGCTCACAAAGGCCAGCCCTCTCTCCCA 987
Qy 72 ----- 72
Db 988 GACTCACCTAGCACTGGTTATTTTGCAGTATGTTCCAAAGACTGGGCGAGCCTTACACCTT 1047
Qy 72 ----- 72
Db 1048 CTGTTTAGCTGGTTCTCTGGCACTGCTCCAGCAGGAAACGGCGCTGTAAGTCCAGGGG 1107
Qy 72 ----- 72
Db 1108 CCTGCACCCACCCCTGCTGCCCATAGCTCAGACGCTCCCGAGGCCAGAGCTGGAGATAAA 1167
Qy 72 ----- 72
Db 1168 CCCAGCTCCAGCTGCTGACTCTGCAGCTGGGCTCTCAGCAGCGGGCTGCTGTGTGT 1227
Qy 72 ----- 72
Db 1228 GGATCTCTGTGTGCCCGGTTCACTGCTGCTTTGTATCTTGTCTCCAGAGCTATCAGTC 1287
Qy 72 ----- 72
Db 1288 ATAGCGGACGGGAGGCTGATGCAGGTTTCCCGGACCTGCAGCATGGAGTGGGGACGG 1347
Qy 73 -----AspAlaCysSer 76
Db 1348 CCGGTGTTCTGTGTGATGCCAGCTGAAGAGCGTGGTGTGTTTACAGATGCTGCTCT 1407
Qy 77 ValProValSerGluIleIleAlaValGluGluThrAspValHisGlyLysHisGlnGly 96
Db 1408 GTGCTGTATCTGAGATCATCGCGTTGAGAAACACAGCTTACCGGGAACATCAAGGC 1467
Qy 97 SerGlyLysTrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArg 116
Db 1468 AGTGGAAATGGCAGAAAATGGAAAAGCCTTACGCTTTTACAGTCTACTGTGTAAGAGA 1527
Qy 117 AlaArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGluLeu 136
Db 1528 GCACGACGGACCGCTGGAAGTGGCGCAGGTACTTCTGGTGTCCAGAGGACAGCTG 1587
Qy 137 CysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLys 156
Db 1588 TGTCACTGTGGCTGCAGACCCCTGCGGGAATGCTGGAGAGCTG----- 1632
Qy 157 HisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGlu 176

Db 1632 ----- 1632
Qy 177 ArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleValThrGlu 196
Db 1633 -----ATTACTGAA 1641
Qy 197 HisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIle 216
Db 1642 CATGCTAATACGCCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATC 1701
Qy 217 ValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyValArgThr 236
Db 1702 GTCTGTGTGGCGGAGATGGTATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGAGC 1761
Qy 237 GlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeu 256
Db 1762 CAGAGGAGCGCGGGGTGCACAGAACACACCCCGGGGTGTGCTGCTCCCGAGTAGCCTC 1821
Qy 257 ArgIleGlyIleIleProAla----- 263
Db 1822 CGGATTGGAATCATTTCCCGCAGGCTTTCGAATGACCTGGCAGGGGACGAGGTGTCTGTC 1881
Qy 263 ----- 263
Db 1882 CTCTCTGGCCCTGTGTCTGGCCCCGAGGGTGGCGGCATGTGTGCACACTTTCACCTCTGTC 1941
Qy 263 ----- 263
Db 1942 ACAGCTCTGGATGTGAGCACCGCAGTCATCCCATTTTATGGATGAAGACAGGAGGACT 2001
Qy 264 -----GlySerThrAspCysVal 269
Db 2002 GGGGAGCATGTGGGCCCCGGTGAGAACGCTGTGTGGCTTGGACGGGTCAACGGACTCGCTG 2061
Qy 270 CysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGly 289
Db 2062 TGTACTCACCGTGGGACACGACGACGAGAACTCGCGCTGCATATCGTGTGGTGG 2121
Qy 290 AspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSer 309
Db 2122 GACTCGCTGGCCATGGATGTGTCTCCTCAGTCACACACACAGACACTCTTCGCTACTCC 2181
Qy 310 ValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysIleArg 329
Db 2182 GTGTCCTCTGGGGTACGGCTTCTACGGGGACATCATCAAGGACAGTCAGAAAGAAACGG 2241
Qy 330 TrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCys 349
Db 2242 TGGTTGGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGACCTTCCTCTCCACCACTGC 2301
Qy 350 TyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArg 369
Db 2302 TATGAAGGGACAGTGTCTCTCTCTCGACACACACAGGTGGATCTCCAAGGATAGG 2361
Qy 370 LysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGlu 389
Db 2362 AAGCCCTGCGGGCAGGATGCTTTGTTTGGAGCAAGCAAGCAGCAGCTGGAGGAGGAG 2421
Qy 390 GlnLysLysAlaLeuTyrGlyLeuGluAlaGluAspValGluGluTrpGlnValVal 409
Db 2422 CAGAAAGAAGCACTGTATGGTTTGAAGCTCGGAGGACGTGGAGGATGGCAAGTCTGTC 2481
Qy 410 CysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerPro 429
Db 2482 TGTGGAGAGTTCTGGCCATCATGCCACAAACATGTCTGTGTGTGCGGAGGCCCC 2541
Qy 430 ArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuArg 449
Db 2542 AGGGGCTCTCCCGGCTGCCACTTGGAGACGGGTCTTCTGACCTCATCTCATCCGG 2601
Qy 450 LysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGln 469

Db 2602 AAATGCTCCAGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACGACGAGACCAG 2661
Qy 470 PheAspPheThrPheValGluValTyrargValLysLysPheGlnPheThrSerLysHis 489
Db 2662 TTTGACTTCATCTTTGTTGAAGTTTATCGCGTCAAGAAATTCAGTTTACGTCGAAGCAC 2721
Qy 490 MetGluAspGluAspSerAspLeuLysGluGlyGlyLysLysArgPheGlyHisIleCys 509
Db 2722 ATGAGAGATGAGACAGCACCTCAAGAGAGGGGGGAGAAAGCGCTTTGGGCACATTTGC 2781
Qy 510 SerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGlu 529
Db 2782 AGCAGCCACCTCTCTGCTGTGCACCGTCTCCACAGCTCTCGAAGCTCGCGGGAG 2841
Qy 530 ValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAla 549
Db 2842 GTCTGTGCACGCCCTGCCATCGAGTCAAGTCCAGTGCACAGTGGTTCGACTCTTTGCA 2901
Qy 550 ArgGlyIleGluGluAsnProLysProAspSerHisSer 562
Db 2902 CGAGGAATTGAAGAGAATCCGAAGCCAGACTCACACAGC 2940
RESULT 11
AAS77728
ID AAS77728 standard; cDNA; 2241 BP.
XX
AC AAS77728;
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13532.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
PT P-PSDB; ABG13541.
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 13532; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: the sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2241 BP; 453 A; 627 C; 700 G; 461 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,03e-184 Length: 2241
Score: 2400.00 Matches: 483
Percent Similarity: 64.39% Conservative: 7
Best Local Similarity: 63.47% Mismatches: 14
Query Match: 79.34% Indels: 258
DB: 5 Gaps: 6

US-10-631-958-11 (1-562) x AAS77728 (1-2241)

```
Qy 6 GlyProAlaProLeuGly-----ValArgAlaProProAlaTTPArgThrSerPro 22
Db 14 GGTTCGCGCCCTATAGACAAGACAGTCACAGGAGCTGGCAGGCTTGGCGACAGCCCG 73
Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTriPVallys 42
Db 74 GCGCGGNGATGGGGCGACGCGGGCGCGAGCGCTGCATCCGTGCTGTGGGTGAAG 133
Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTfPTrPArgSer 62
Db 134 CAGCAGCGCTGCGCGCTGAGCTGGAGCCGCGCGGCTCTGCTGCTGTGGTGGCGAGC 193
Qy 63 ProGlyProGlyAlaGlyAlaProGlyAla----- 72
Db 194 CCGGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 252
Qy 72 ----- 72
Db 253 GACTCACCTAGCAGTGGTTATTTTGCAGTATGTTTCCAGACTGGGCGGCTTTACACCTT 312
Qy 72 ----- 72
Db 313 CTGTTTAGCTGTTTCTCTGCACTGTCTCCAGCAGGGAACGGCGCTGGTACTGCCAGGG 372
Qy 72 ----- 72
Db 373 CCTGCACCCACCTGCTGCCCTCCATAGCTCAGACGTCCTCCGAGGCCAGAGCTGGAGATAA 432
Qy 72 ----- 72
Db 433 CCCAGCTCCCACTGCTGACTCTGCACGCTGGCCCTCTCCAGCCAGCGGGCTGTGCTGT 492
Qy 72 ----- 72
Db 493 GGATCTCTGGTGTCCCGGTTCACTGCGCTGTTCTTGTATCTTGTCCAGAGCTATCAGTC 552
Qy 72 ----- 72
Db 553 ATAGCGGACCGGAGGCTGATGCAGGTTTCCCGCACTGTACAGCATGGAGTGGGACGG 612
Qy 73 ----- AspAlaCysSer 76
Db 613 CGGGTGTTCTGTGTGATGCCAGCTGAAGGACGTGGTGTGTTTACAGATGCTGTCTCT 672
Qy 77 ValProValSerGluIleAlaValGluThrAspValHisGlyLysHisGlnGly 96
Db 673 GTGCTGTATCTGAGATCATCGCGTTGAGGAAACAGAGCTTTCACGGGAAACATCAAGGC 732
Qy 97 SerGlyLysTrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArg 116
Db 733 AGTGGAAATGGCAGAAATGGAAAGCCTTTACGCTTTTACAGTTTCACTGTGTGAAGAGA 792
```

```
Qy 117 AlaArgArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeu 136
Db 793 GCACGACGGCACCGCTGGAGTGGCGCAGGTGACTTCTTGGTGTCCAGAGGAGCAGCTG 852
Qy 137 CysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLys 156
Db 853 TGTCACTTGTGGTGCAGACCCCTGCGGGAGATGCTGGAGAAGCTG----- 897
Qy 157 HisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGlu 176
Db 897 ----- 897
Qy 177 ArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGlu 196
Db 898 -----ATTACTGAA 906
Qy 197 HisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIle 216
Db 907 CATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGCGGCATC 966
Qy 217 ValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThr 236
Db 967 GTCTGTGTCGCGGAGATGGTAIGTTTCAGCGAGTGTCTGACGGTCTGATTTGGAGAGCG 1026
Qy 237 GlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeu 256
Db 1027 CAGAGGAGCGCGGGTTCGACACAGAACCCCGGGCTGTGTGGTCCCCCAGTAGCCTC 1086
Qy 257 ArgIleGlyIleIleProAla----- 263
Db 1087 CGGATTGGAATCATTCCTCCGACAGCATGTGGGCGCCGCGTGAGAACCGCTGGTGGCTTGGAC 1146
Qy 264 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 283
Db 1147 GGGTCAACGAGACTGCGTGTGTACTCCACCCTGGGACACGAGCAGCAGAAACCTCGCG 1206
Qy 284 LeuHisIleValVal----- 288
Db 1207 CTGCATATCGTTGTGGCTGTCTGCCGAGGCTCGGAAACCCCGCATCCCGGCATACT 1266
Qy 288 ----- 288
Db 1267 GCCTGTGGCAGTGGGCACGGGCAGCTGTGTCTGGCTGCCAGCTGTGGAACGCATGCCTC 1326
Qy 289 -----GlyAspSerLeuAlaMetAsp 295
Db 1327 TGTGAGCCCTCGAGGCTTCAGTCCAGATCCAGAGCCCCCGGGAGCTCGCTGGCCATCGAT 1386
Qy 296 ValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyr 315
Db 1387 GTGTCTCTCAGTCCACCAACACAGCACACTCTTTCGCTACTCCGTGTCCCTGCTGGGCTAC 1446
Qy 316 GlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArg 335
Db 1447 GGCTTCTACCGGGACATCATCAAGGACAGTGAAGAAACCGTGGTGGGTCTTGGCAGA 1506
Qy 336 TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSer 355
Db 1507 TACGACITTTTCAAGTATAAGACCTTCTCTCCACCACTGTCTATGAAGGACAGTGTCC 1566
Qy 356 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 375
Db 1567 TTCTCTCTCCACAAACACACCGTGGGATCTCCAAGGAATAGGAAGCCCTGCGGGCAGGA 1626
Qy 376 CysPheValCysArgGlnSerLysGlnLeuGluGlnLysLysAlaLeuTyr 395
Db 1627 TGCTTTTTCAGGCAAAACAGCAGCAGCTGAGGAGGAGGAGGAGAAAGCACTGTAT 1686
Qy 396 GlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAla 415
Db 1687 GGTTTGGAAGCTCGGAGGAGCAGTGGAGAGTGGCAAGTCTGTCTGTGGGAAGTTTCTGCC 1746
Qy 416 IleAsnAlaThrAsnMetSerCysAlaCysArgSerProArgGlyLeuSerProAla 435
```

Db	1747		ATCAATGCCAACAATGTCCTGCTGTGCGCGGAGCCCAAGGGGCTCTCCCGGCT	1806
Qy	436		AlaHisLeuGlyAspGlySerSerAspLeuileLeuileArgLysCysSerArgPheAsn	455
Db	1807		GCCACCTGGGAGACGGGTCTTCGACCTCATCTCTCATCCGGAATGCTCAGGTTCAAT	1866
Qy	456		PheLeuArgPheLeuileArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheVal	475
Db	1867		TTTCTGAGATTCTCATCAGGCACACCAACCAGCAGCAGCTTTGACTTCACTTTTGT	1926
Qy	476		GluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSer	495
Db	1927		GAAGTTTATCGCGTCAAGAATTCTCTGTTTACGTCGAAGCACATGGAGGATGAGGACAGC	1986
Qy	496		AspLeuLysGluGlyGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCys	515
Db	1987		GACCTCAAGAGAGGGGGGAAGAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCTGTC	2046
Qy	516		CysCysThrValSerAsnSerSerTyrAsnCysAspGlyGluValLeuHisSerProAla	535
Db	2047		TGCTGCACCGTCTCCACAGCTCTGGAACCTGCAGCGGGAGGTCTCTGCACAGCCCTGCC	2106
Qy	536		IleGluValArgValHisCysGlnLeuValArgLeuPheIleArgGlyIleGluGluAsn	555
Db	2107		ATCAGGTCAGG-----GCGGTGGTGATGGCAAGAGGCTGTGCAAAAGGTC	2151
Qy	556	Pro	556	
Db	2152	CCA	2154	
RESULT 12				
ACC70838				
ID	ACC70838	standard; DNA; 3975 BP.		
XX	XX			
AC	ACC70838;			
XX	XX			
XX	20-NOV-2003	(first entry)		
XX	Human Sphingosine kinase 4 coding sequence.			
DE	DE			
XX	XX			
KW	Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;			
KW	platelet transfusion; platelet stabiliser; gene; ds.			
XX	XX			
OS	Homo sapiens.			
XX	XX			
Key	Location/Qualifiers			
CDS	1..1251			
FT	/*tag= a			
FT	/product= "Sphingosine kinase 4"			
XX	XX			
PN	WO2003031627-A1.			
XX	XX			
PD	17-APR-2003.			
XX	XX			
PF	28-SEP-2001; 2001WO-JP008537.			
XX	XX			
PR	28-SEP-2001; 2001WO-JP008537.			
XX	XX			
PA	(HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.			
PA	(CHBI-) CHEM BIOLOGY INST.			
XX	XX			
PI	Igarashi Y, Kihara A;			
XX	XX			
DR	WPI; 2003-354917/33.			
DR	P-PSDB; ABR56301.			
XX	XX			
PT	Platelet derived polypeptides with sphingosine kinase activity for			
PT	treatment of sphingosine related disorders.			
XX	XX			
PS	Claim 4; Page 30-31; 39pp; Japanese.			
XX	XX			
CC	The present sequence is the coding sequence for human sphingosine kinase			

CC	4	(SPHK4). The kinase can be used for the diagnosis and treatment of sphingosine related disorders. The kinase can also be potentially used for controlling toxicity of platelet transfusion and as a platelet stabiliser	XX
CC	Sequence	3975 BP; 943 A; 978 C; 1042 G; 1012 T; 0 U; 0 Other;	XX
CC	Alignment Scores:		XX
CC	Pred. No.:	1.07e-168 Length: 3975	XX
CC	Score:	2210.00 Matches: 416	XX
CC	Percent Similarity:	100.00% Conservative: 0	XX
CC	Best Local Similarity:	100.00% Mismatches: 0	XX
CC	Query Match:	73.06% Indels: 0	XX
CC	DB:	8 Gaps: 0	XX
CC	US-10-631-958-11	(1-562) x ACC70838 (1-3975)	XX
Qy	147	MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly	166
Db	1	ATGCTTGAGAAAGCTGAGCGTCAGACCAAGCATTTACTTGGTATTTATCAACCGTTTGA	60
Qy	167	GlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAla	186
Db	61	GGAAAGGACAAAGCAAGCGGATATGAAAGAAAGTGGCACCACTGTTCACTTAGCC	120
Qy	187	SerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyr	206
Db	121	TCCATCACCACCTGACATCATCTGTTACTGAAACATGCTAATAGGCCCAAGGAGACTCTGTAT	180
Qy	207	GluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSer	226
Db	181	GAGATTAAACATAGCAAAATACGACGGCATCGTCTGTGCGCGGAGATGTTATGTTCCAG	240
Qy	227	GluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHis	246
Db	241	GAGGTGTCGACGCTCTGATTGGGAGGACGACAGGAGCGCGGGGTTCGACCAGAACCAAC	300
Qy	247	ProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThr	266
Db	301	CCCCGGGCTGTGCTGGTCCCAGTAGCCTCCGGATTTGGAATCATTTCCGCGAGGGTCAACG	360
Qy	267	AspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIle	286
Db	361	GACTGCGTGTTTACTCCACGTTGGGCACACGACGACGAGAAACCTCGCGCGTGCATATC	420
Qy	287	ValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeu	306
Db	421	GTGTGTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCACAAACAGCACACTCCTT	480
Qy	307	ArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleIysAspSerGlu	326
Db	481	CGCTACTCCGTTGCCCTGTGGCTACGGCTTCTACCGGGACATCATCAAGGACAGTGG	540
Qy	327	LysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSer	346
Db	541	AAGAAACGGTGTGGTCTTGGCCATACACACTTTTCAGGTTTAAAGACCTTCTCTCCTCC	600
Qy	347	HisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerPro	366
Db	601	CACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACACGCGTGGGATCTCCA	660
Qy	367	ArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeu	386
Db	661	AGGANTAGGAAGCCCTCCCGGGCAGGATGCTTTGTGTCAGGCAAGCAAGCAGCAGCTG	720
Qy	387	GluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaIaGluAspValGluGluTrp	406
Db	721	GAGAGAGAGCAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGCACTGGAGAGTGG	780
Qy	407	GlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArg	426
Db	781	CAAGTCGTCTGTGGGAAGTTCTGGCCATCAATGCCACAAACATGCTCTGCTGTTGTGCG	840

The present sequence is the coding sequence for human sphingosine kinase

Qy 427 ArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLe 446
 Db 841 CGAGACCCAGGGGCGCTCTCCCGCGCTGCCACTTGGAGAGCGGCTCTTGACCTCATC 900
 Qy 447 LeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuArgHisThrAsnGln 466
 Db 901 CTATCCGGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACCAACACAG 960
 Qy 467 GlnAspGlnPheAspPheThrPheValGluValThrArgValLysLysPheGlnPheThr 486
 Db 961 CAGGACCATTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAATTTCCAGTTTACG 1020
 Qy 487 SerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLysLysArgPheGly 506
 Db 1021 TCGAAGCACATGGAGGATGAGGACAGCGACTCAAGGAGGGGGGAGAGCGCTTTGGG 1080
 Qy 507 HisLeuCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCys 526
 Db 1081 CACATTTGCAGCAGCCACCCCTCTGCTGTGCAACCGTCTCCAAACAGCTCTGGAACAGC 1140
 Qy 527 AspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArg 546
 Db 1141 GACGGGAGGTCCTGCAACAGCCCTGCCATCGAGGTCAAGTCCACTGCCAGCTGGTTGCA 1200
 Qy 547 LeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 562
 Db 1201 CTCTTTGCAGGAGAAATTGAAGAGATCCGAGGACGACTCACACAGC 1248

RESULT 13

AAS77730

ID AAS77730 standard; cDNA; 2186 BP.

XX AAS77730;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #13534.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG13543.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensic, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity.

XX Claim 1; SEQ ID NO 13534; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX sequences. (I) is useful as hybridisation probes, polymerase chain

XX reaction (PCR) primers, oligomers and for chromosome and gene mapping,

XX and in recombinant production of (II). The polynucleotides are also used

XX in diagnostics as expressed sequence tags for identifying expressed

XX genes. (II) is useful in gene therapy techniques to restore normal

XX activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensic, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: the sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 2186 BP; 455 A; 637 C; 669 G; 425 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.81e-129 Length: 2186
 Score: 1714.50 Matches: 345
 Percent Similarity: 63.65% Conservative: 0
 Best Local Similarity: 63.65% Mismatches: 0
 Query Match: 56.68% Indels: 197
 DB: 5 Gaps: 3

US-10-631-958-11 (1-562) x AAS77730 (1-2186)

Qy 195 ThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluLeuLeuAsnIleAspLysTyrAsp 214

Db 3 ACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGAC 62

Qy 215 Gly----- 215

Db 63 GGATGTACTACTGCCCATTCAGAGTGCCTGTCTCGAAACACACAGCCCGCTGGGAG 122

Qy 215 ----- 215

Db 123 GCTGCCCGGTGCAAGGTGTAGGCTACGGGGAAGGCAAGAGCCTTCCCATGGAGCCG 182

Qy 215 ----- 215

Db 183 GGCCCAAGCAGATGCTGCCAAGGGCCAGGTGTGACAGGCTGCCCTCTCTGAAGCTGGAA 242

Qy 215 ----- 215

Db 243 GCCTCAGGCTGTGCGCTCAGAGGGCCGACCTGTCAGAGGCCCTGACCGGCTGTGTTC 302

Qy 216 -----IleValCysValGlyAspGlyMetPheSerGluValLeuHisGlyLeu 232

Db 303 TCCTGCAGCATCGTCTGTGTCGGCGAGATGGTATGTTTCAGCGAGGTGCTGCACGGTCTG 362

Qy 233 IleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuVal 252

Db 363 ATTGGGAGGACGACAGAGCGCGGGTGCAGCAACACCACCCCGGGTGTGTCTGGTC 422

Qy 253 ProSerSerLeuArgIleGlyIleIleProAla----- 263

Db 423 CCCAGTAGCCTCCGATTTGAATCATTTCCCGCAGGCTTTGCAATGACCTGGCAGGGGAC 482

Qy 263 ----- 263

Db 483 GAGGTGTCTGTCTCTCTGCGCCTGTGTGTGCGCCCGAGGGTGGCGGCATGTTGCACACT 542

Qy 263 ----- 263

Db 543 TTCACCTCTGTCACAGCTCTGGGATGTGAGACCGCAGTCATCCCATTTTATGGATGAA 602

Qy 264 -----GlySer 265

Db 603 GACAGGAGGACTGGGAGCATGTGGGCCCGGTGAGAACGCTGTGTGGCTTGGACGGGTCA 662

Qy 266 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 285


```
Db 663 ACGGACTGCGTGTGTACTCCACCGTGGCGACCGAGCGCAGAAACCTCGGCGCTGCAT 722
Qy 286 lleValVal----- 288
Db 723 ATTCGTTGTTGGCTGTGCCCCGAGGCTCGAAACACCCGCGTCCGCCATACATCGCTGT 782
Qy 288 ----- 288
Db 783 GCGAGTGGGCACGGCAGCTGTGCTGGCTGCCAGCTGTGGAACGCATGCTCTGTGTGAG 842
Qy 289 -----GlyAspSerLeuAlaMetAspValSer 297
Db 843 GCGTCGAGGCTTCAGTCCAGGATGCAGAGCCCCGGGACTCGCTGCACATGATGTGCC 902
Qy 298 SerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 317
Db 903 TCAGTCCACCAACACAGCACACTCCCTCGCTACTCCGTGTCCTGCTGGCTTACGGCTC 962
Qy 318 TyrGlyAspIleIleValAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAsp 337
Db 963 TAGGGGACATCATCAAGGACAGTGAAGAAACCGTGTGGTCTTGGCCAGATACGAC 1022
Qy 338 PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu 357
Db 1023 TTTTCAGGTTAAAGACCTCTCTCCACCACTGCTATGAAGGCACAGTGTCTCTCTC 1082
Qy 358 ProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 377
Db 1083 CCTGCACAACACACCGTGGGATCTCCAAGGGATAGGAAGCCCTGCGGCGAGGATGCTTT 1142
Qy 378 ValCysArgGlnSerLysGlnLeuGluGluGlnLysLysLysLysLysLysLysLys 397
Db 1143 GTTTCGAGCAAGCAAGCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTGGTTG 1202
Qy 398 GluAlaLysGluAspValGluGlnValValCysGlyLysPheLeuAlaIleAsn 417
Db 1203 GAAGTCGCGAGGACGTGGAGGATGGCAAGTCTGTGGGAAGTTCTGGCCATCAAT 1262
Qy 418 AlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHis 437
Db 1263 GCCAACAACATGTCCTGTGCTGTGCGCGAGCCCGAGGGGCTCTCCCGCTGCCAC 1322
Qy 438 LeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeu 457
Db 1323 TTGGAGACGGGTCTCTGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTG 1382
Qy 458 ArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluVal 477
Db 1383 AGATTTCTCATCAGGCACACCAACACGAGGACCACTTTGACTTCCTTTTGTGAAGTT 1442
Qy 478 TyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu 497
Db 1443 TATCGCTCAAGAAATTCAGTTTACGTCGAAGCACATGGAGGATGAGGACAGCGACCTC 1502
Qy 498 LysGluGlyGlyLysLysArgPheGlyHisLysCysSerSerHisProSerCysCysCys 517
Db 1503 AAGGAGGGGGGAAAGCGCTTTTGGGCACATTTGCGACAGCCACCCCTCTGCTGCTGC 1562
Qy 518 ThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGlu 537
Db 1563 ACCGTCTCCACAGCTCTCTGGNACTGCCAGCGGGGAGGTCTTGCCACAGCCCTGCCATC 1622
Qy 538 ValArg 539
Db 1623 GTCAGA 1628
RESULT 14
ID ABL40822
XX standard; cDNA; 979 BP.
AC ABL40822;
XX
DT 03-JUL-2002 (first entry)
```

```
XX Human sphingosine kinase-like protein encoding cDNA.
DE
XX
KW Human sphingosine kinase-like protein; intracellular signalling; gene;
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
CDS 2..789
/*tag= a
/product= "sphingosine kinase-like protein"
/note= "start and stop codons are not indicated"
XX
WO200228906-A2.
XX
11-APR-2002.
XX
05-OCT-2001; 2001WO-EP011516.
XX
06-OCT-2000; 2000US-0238005P.
XX
23-AUG-2001; 2001US-0314113P.
XX
(PARB ) BAYER AG.
XX
Kossida S, Encinas J;
XX
WPI; 2002-340094/37.
XX
P-PSDB; ABB07854.
XX
New reagent for modulating the activity of sphingosine kinase-like
XX protein polypeptide or polynucleotide and treating cancer, asthma,
XX allergy, an autoimmune disease, or a central or peripheral nervous system
XX disorder.
XX
Claim 1; Fig 1; 120pp; English.
XX
The invention relates to a human sphingosine kinase-like protein. The
XX polypeptide can be expressed by standard recombinant methodology. The
XX sphingosine kinase-like protein and gene can be used to regulate
XX intracellular signalling and consequently cell proliferation and
XX apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
XX asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
XX peripheral nervous system disorders (e.g. Parkinson's disease). The
XX present sequence represents the human sphingosine kinase-like protein
XX encoding cDNA
XX
SQ Sequence 979 BP; 237 A; 259 C; 271 G; 212 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.32e-123 Length: 979
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 54.23% Indels: 11
DB: 6 Gaps: 1

US-10-631-958-11 (1-562) x ABL40822 (1-979)
Qy 155 ProlysHisLeuLeuValPheIleAsnProPheGlyLysGlyGlnGlyArgIle 174
Db 2 CCAAGCATTTTACTGGTATTATTCACCCGCTTTGGAGGAAAGGACAGGCGGATA 61
Qy 175 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle--- 193
Db 62 TATGAAAGAAAGTGGCACCACCTGTTCCCTTACCTCCATCCACTGACATCATCGGT 121
Qy 194 -----ValThrGluHisAlaAsnGlnAlaLysGlu 203
Db 122 AACAAATTCATGTTAACTATGTAGAAAGTAATTACTGAACATGCTAATCAGGCCAGGAG 181
Qy 204 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 223
```


This Page Blank (uspto)

Result No.	Score	Query			DB	ID	Description
		Match	Length	*			
1	3025	100.0	4413	6	AX457006	Sequence	
2	3012	99.6	2042	9	HA545788	Homo sapi	
3	3012	99.6	4445	9	AB079066	Homo sapi	
4	3004	99.3	4463	6	BD102675	Ceramide	

TFWCPBQLCHLWLTQIREMLEKLTSRPKHLVFINPFGKGQKRIYERKVAFLPTL
ASITTDIIVTEHANOAKETLEYINIDRYDGIIVCGDGMFSEVLHGLIGRTORSAGVD
QNHPRVLVSRLRIGIIPAGSTDCVCTVGTSDAETSAHLHIVGDSLAMDVSSVHH
NSTLLRYSVLGLGVFGYDIIKDKSEKKRWLGARYDFSGLKTFLSHHCVGTVSFLPA
QHTVSPDRKPCRAGCFVCRSQKQLEBEOKKALYGLEAAEDVEEHQVYCGKFLAIN
QTNMSCARSPRGLSPAHLHGDSLLILRKSRFNFLRLIRHTNQQDQDFTFV
EYVRVKFQTSKMEDEDEDLKEGGKKRFGHICSSHPSCCCTVSNMNCDEVLHS
PAIEVRVHCQLVRLFARGIBENPKPDGSHS*

ORIGIN

Alignment Scores:

Pred. No.: 3.33e-199 Length: 4413
Score: 3025.00 Matches: 562
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-631-958-11 (1-562) x AX457006 (1-4413)

Qy	1	HisGluAlaAAsnGlyProAlaProLeuGlyValArgAlaProProAlaTtpArgThr	20
Db	1	CACAGGCGGTAAAGCTCGCGGCCCTCGCGCTCGCGGCCCCAGCGCTGGCGAGC	60
Qy	21	SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrp	40
Db	61	AGCCCGCGCGGAGATGGGGCGCACGGGGCGCGGAGCGGCTGCAATCGTCTGTGG	120
Qy	41	VallysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrp	60
Db	121	GTGAAGCAGCAGCGCTCGCGCTGAGCTTGAGCGCCGCGGGCTCTGCTCGCTGTGG	180
Qy	61	ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSer	80
Db	181	CGGAGCCCGGGGCCGCGGAGCCGCCGCCCGCGCGGATGCTCTGTGCGCTGTATCT	240
Qy	81	GluIleLeAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp	100
Db	241	GAGATCATCGCGTTGAGGAAACAGAGCTTCACGGGAAACATCAAGCAGTGGGAAATGG	300
Qy	101	GlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHis	120
Db	301	CAGAAATGGGAAAGCCCTACGCTTTACAGCTTCACTGTGTGTAAGAGCAGCAGCGCAC	360
Qy	121	ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrp	140
Db	361	CGCTGGAAAGTGGGCGCAGTGACTTTCTGGTGTCCAGAGGACAGCTGTCTACTGTGG	420
Qy	141	LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal	160
Db	421	CTGCAGACCCCTGCGGAGATGCTGGAGAAAGCTGACGCTCCAGACCAAGCATTTACTGTA	480
Qy	161	PheIleAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArgLysValAla	180
Db	481	TTTATCAACCCGTTTGGAGGAAAGGACCAAGCGGATATATGAAAGGAAAGTGGCA	540
Qy	181	ProLeuPheThrLeuAlaSerIleThrThrAspIleLeuValThrGluHisAlaAsnGln	200
Db	541	CCACTGTTCCTTAGCCTTCATCACCACTGACATCATCGTTACTGAACTGCTTAATCAG	600
Qy	201	AlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly	220
Db	601	GCCAAGGAGACTCTGTATGAGATTAAACATACAAAATACGACGCGCATCTGTGTGCGC	660
Qy	221	GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla	240
Db	661	GGAGATGGTATGTTTACGAGAGTGTGTCACCGTCTGATTTGGAGGAGCGCAGAGCGCC	720
Qy	241	GlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle	260
Db	721	GGGGTCGACCAAGAACCCCGGGCTGTGTGGTCCCGAGTAGCCTCCGGATTGGATC	780
Qy	261	IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu	280

Db	781	ATTCCCGCAGGGTCAACGGACTCGGTGTGTATTCCACCGTGGGCACACGCGCAGAA	840
Qy	281	ThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHis	300
Db	841	ACCTCGCGCTGCATATCGTGTGTGGGACTCCGCTGCCCATGATGTCTCTAGTCCAC	900
Qy	301	HisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp	320
Db	901	CACAACAGCACACTCCTTCGTACTCCGTCTCCTGCTGGCTTACGGCTTCTACGGGAC	960
Qy	321	IleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGly	340
Db	961	ATCATCAAGGACAGTGAAGAAACGGTGGTGGTCTTGGCTTCCAGATACGACTTTTCAGGT	1020
Qy	341	LeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGln	360
Db	1021	TTAAAGACCTTCTCTCCACCACCTGCTATGAAGGAGCAGTGTCTTCTCCCTCGACAA	1080
Qy	361	HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg	380
Db	1081	CACACGGTGGGATCTCCAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGCAGG	1140
Qy	381	GlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAla	400
Db	1141	CAAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAACACTGTATGTGTTTGGAGCTGCG	1200
Qy	401	GluAspValGluLutrpGlnValValCysGlyLysPheLeuAlaIleAenAlaThrAsn	420
Db	1201	GAGACGTGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAAC	1260
Qy	421	MetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp	440
Db	1261	ATGTCCTGTCTTGTGCGCGAGGCCCCAGGGGCTCTCCCCGCTGCCCATTTGGGAGAC	1320
Qy	441	GlySerSerAspLeuIleLeuLysCysSerArgPheAsnPheLeuArgPheLeu	460
Db	1321	GGGTCTTCTGACCTCATCTCATCCGGAATGCTCCAGGTTCATTTTCTGAGATTTCTC	1380
Qy	461	IleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluValTyrArgVal	480
Db	1381	ATCAGGCACACCAACACAGCAGGACCACTTGTGACTTCACTTTGTTGAAGTTTATCGGTC	1440
Qy	481	LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly	500
Db	1441	AAGAAATTCAGTTTACGTCGAAGCACATGAGGATGAGGACAGCGACCTCAAGAGGGG	1500
Qy	501	GlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSer	520
Db	1501	GGGAAGAAGCGCTTGGGCACATTTGCAGCAGCACCCCTCTGCTGCTGCTGCTGCTCC	1560
Qy	521	AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVal	540
Db	1561	AAAGCTCTCGGAACCTCGACGGGAGGCTCTGCACAGCCCTGCCATCGAGGTCAGAGTC	1620
Qy	541	HisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSer	560
Db	1621	CACCTGCCAGCTGTTGACTCTTTTGACGAGGAATTGAAGAGAAATCCGAAGCCAGACTCA	1680
Qy	561	HisSer 562	
Db	1681	CACAGC 1686	

RESULT 2

HSA457828 2042 bp mRNA linear PRI 19-APR-2002
LOCUS HSA457828
DEFINITION Homo sapiens mRNA for putative lipid kinase (LK4 gene).
ACCESSION AJ457828
VERSION AJ457828.1 GI:20269072
KEYWORDS lipid kinase; LK4 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Van Veldhoven, P.P.
A search for lipid kinases
Unpublished
2 (bases 1 to 2042)
Van Veldhoven, P.P.
Direct Submission
Submitted (18-APR-2002) Van Veldhoven P.P., Farmakologie,
K.U.Leuven, Herestraat, B-3000 Leuven, BELGIUM
Location/Qualifiers
1. 2042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="22"
/map="22q13"
1. 2042
/gene="LK4"
71..1684
/gene="LK4"
/codon_start=1
/product="putative lipid kinase"
/protein_id="CAD29884.1"
/db_xref="GI:20269073"
/db_xref="GOA:Q8TC70"
/translation="MGATCAAEPLQSLVWVQQRCAVSLPARALWRRSPGPCAGA
PGADACSVPSVEI IAVETDVHGKHGSGKQMEKPYAFTVCHVCRARRRHWKAOV
TFWCPEROLCHLWLOTLEMLEKLTSPKHLVFINPFGKGQKRIYVERKVAFL
ASITDIIVTEHANQAKETLYEINLDKVDGLVCGDGMFSEVLHGLIGRTQRSAGVD
QNHPRVLPSSRLIGIIPAGSTDVCYSTVSTGSDAETSALHI VVGDSLAMDVSVHH
NSTLLRYSVLGYFGDIIDKSEKKRWGLARYDFSLKTFILSHHCYEGTVSFLPA
QHTVGSPRDRKPRACGFCVRSQKQLEEKQKALYGLEAAEDVEEVMQVCGKFLAIN
ATNMSCACRRSPRGLSPAHLGDGSSDLILRKCSRNFRLRLRHTNQOQDFTFV
EYVRVKFQFTSKHWEDESDLKEGGKRFHGHICSSHPSCCCTVSNSSNWCNCGEVLHS
PAIEVRVHCQLVRLFARGIEENPKPDSHS"

ORIGIN

Alignment Scores:
Pred. No.: 1e-198 Length: 2042
Score: 3012.00 Matches: 560
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.57% Indels: 0
DB: Gaps: 0

US-10-631-958-11 (1-562) x HSA457828 (1-2042)

QY 3 AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerPro 22
DB 2 GCCGCTAACGGTCCGGCGCCCTCGCGCGTCCGCGCGCCCGCCAGCCTGGCGGACGACCGCG 61
QY 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpVallys 42
DB 62 CGCGCGGAGATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 121
QY 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpArgSer 62
DB 122 CAGCAGCGCTGCGCGCTGAGCCTGGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181
QY 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle 82
DB 182 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 241
QY 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys 102
DB 242 ATCGCGCGTTGAGGAAACACAGCTTCACGGGAAACATCAAGGCAGTGGGAAATGGCAGAAA 301
QY 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrp 122
DB 302 ATGGAAAGACCTTACGCTTTTACAGTTTCACTGTGTATAGAGAGACGACGACCGACCGCTGG 361

QY 123 LysTrpAlaGlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGln 142
DB 362 AAGTGGCGCGCAGGTGACTTTCTGGTGTCCAGAGAGCAGCTGTGTCACTTGTGGCTGCAG 421
QY 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162
DB 422 ACCCTGCGGAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTACTGGTATTATC 481
QY 163 AsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu 182
DB 482 AACCCGTTGGAGGAAAGGACAGGCAAGCGGATATATGAAGAAAAAGTGGCACCACCTG 541
QY 183 PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys 202
DB 542 TTCACCTTAGCTCCATCACCACATGACATCATCTGTACTTGAACATGCTTAATCAGGCCAAG 601
QY 203 GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp 222
DB 602 GAGACTCTGTATGAGATTACATAGACAAATACACCGGCATCGTCTGTGTGGCGGAGAT 661
QY 223 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 242
DB 662 GGTATGTTACGCGAGGTGCTGACCGTCTGATTGGGAGGACGAGAGCGCGCGGCTC 721
QY 243 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIlePro 262
DB 722 GACCAGAACCAACCCCGCGGTGCTGTGTCCTCCAGTAGCTCCGATTTGGAAATCATTCCTC 781
QY 263 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282
DB 782 GCAGGTCACCGACTGCGTGTGTATTCTCCACCGTGGCCACGAGGACGAGAACCTCG 841
QY 283 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn 302
DB 842 CGCTGTCATATCGTTGTTGGGACTCGCTGGCCCATGATGTGTCTCAGTCCACCAAC 901
QY 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322
DB 902 AGCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 961
QY 323 LysAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342
DB 962 AAGGACAGTGAAGAAGAAACGGTGGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAG 1021
QY 343 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 362
DB 1022 ACCTTCTCTCTCCACCATGCTATGAAGGACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1081
QY 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382
DB 1082 GTGGGATCTCCAAAGGATAGGAAGCCCTGCGCGGACGAGATGCTTTGTTTGGAGGCAAGC 1141
QY 383 LysGlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402
DB 1142 AAGCAGCAGCTGGAGGAGGACAGAGAACCATGTATGTGTTTGGAGCTGCGGAGGAC 1201
QY 403 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 422
DB 1202 GTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTC 1261
QY 423 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442
DB 1262 TGTGCTGTCTCCCGGAGCCCGGAGGCGCTCTCTCCCGGCTGCGCCACTTGGGAGAGCGGGTCT 1321
QY 443 SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg 462
DB 1322 TCTGACCTCATCTCATCCGGAATGTCTCAGGTTCAATTTTCTGAGATTCTCATCAGG 1381
QY 463 HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482
DB 1382 CACACCAACACGAGGACCGTGTGACTTCTCACTTTTGTGAAGTTTATCGCGTCAAGAAA 1441
QY 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys 502

Db	1442	TTCAGATTTTACGTCGAAGCATCGAGGATGAGACGACGACCTTCAAGGAGGGGGGAAG	1501
Qy	503	LysArgPheGlyHisIleCysSerHisProSerCysCysCysThrValSerAsnSer	522
Db	1502	AAGCGCTTTGGGCAATTTGACGAGCACCCTTCCTGCTGCTGCACCGTCTCCACAGC	1561
Qy	523	SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys	542
Db	1562	TCCTGGAACTGCGACGGGGAGTCTTCACACAGCCCTGCCATCGAGGTCCAGATCCACTGC	1621
Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer	562
Db	1622	CAGCTGGTTCGACTCTTTGACGAGGAATTCAGAGAATCCGAAGCCAGACTCCACAGC	1681
RESULT 3			
LOCUS	AB079066	4445 bp mRNA linear	PRI 27-JUN-2002
DEFINITION	Homo sapiens cerk mRNA for ceramide kinase, complete cds.		
ACCESSION	AB079066		
VERSION	AB079066.1	GI:21624339	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Sugiura,M., Kono,K., Liu,H., Shimizugawa,T., Minekura,H., Spiegel,S., and Kohama,T.		
TITLE	Ceramide Kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND FUNCTIONAL CHARACTERIZATION		
JOURNAL	J. Biol. Chem. 277 (26), 23294-23300 (2002)		
PUBMED	11956206		
REFERENCE	2 (bases 1 to 4445)		
AUTHORS	Sugiura,M., Kono,K., Shimizugawa,T., Minekura,H., Spiegel,S. and Kohama,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-JAN-2002) Masako Sugiura, Sankyo Co., Ltd., Pharmacology and Molecular Biology Research Laboratories, 2-58 Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan (E-mail:msugiura@shina.sankyo.co.jp, Tel:81-3-3492-3131, Fax:81-3-5436-8565)		
FEATURES	Location/Qualifiers		
source	1..4445		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
gene	1..4445		
	/gene="cerk"		
CDS	124..1737		
	/gene="cerk"		
	/codon_start=1		
	/product="ceramide kinase"		
	/protein_id="BAC01154.1"		
	/db_xref="GI:21624340"		
	/translation="MGATGAAPLQSVLWVKQRCVASLEPARALLWRSPGPGAGA PGADACSPVSEIITAVETDVHGKQSGKWKPKYAFVCHVKARHRKWKQVLTWCPERQLQHQLQTLREMLEKTSRPHLLVFINPPGKGQGRKLYERKVALPTL ASITDILIVHEHANQAKETIYEINRIGIDGIVCVGGCMFSEVHLIGRTQKSGVD QNHPRVLVPSSRGRIIPAGSTDCCVYSTGTSDAETSAHLHIVGDSLAMDVSSVHH NSTLLYSVLGIGFYGDIIKQSEKRWGLGARYDFSGSLKTFLSHHCYEGTYSFLPA QHTVGSRDPRKPCAGFCVCRQQLKEEKALGLEAAEDVEEWVVCVGKFLAIN ATNNSCACRRSPRLGSPAHLGDGSDLLIRKCSRFNLFRLIRHTNQODQDFDTFV EVYRVKFPQTSKHMEDSDLEKGGKRGFKHCSSHPSCCTVSSNWNCDGEVLHS FAIEVRVHCQVLRLFARFGENPKPDSHS"		
ORIGIN			
Alignment Scores:			
Pred No.:	2,66e-198	Length:	4445
Score:	3012.00	Matches:	560
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.57%	Indels:	0

Db	1075	ACCTTCTCTCCACACACTGCTATGAGGACAGTGTCTTCTCTCCCTGCACACACAG	1134
Qy	363	ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer	382
Db	1135	GTGGGATCTCCAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTGTGTTCAGGCAAGC	1194
Qy	383	LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrglyLeuGluAlaAlaGluasp	402
Db	1195	AAGCAGCAGCTGAGAGAGAGAGAGAGAACACCTGTATGGTTTGAAGCTGCGAGAGAC	1254
Qy	403	ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer	422
Db	1255	GTGGAGGAGTGCAGAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC	1314
Qy	423	CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer	442
Db	1315	TGTGCTTGTGCGCGAGGCCACAGGGGCTCTCCCCGGCTGCCACTTGGGAGCGGTCT	1374
Qy	443	SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg	462
Db	1375	TCTGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG	1434
Qy	463	HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyArgValLysLys	482
Db	1435	CACACCAACGACGACGACGATTTGACTTCTCTTTGTGTGAAGTTATCGCGTCAAGAAA	1494
Qy	483	PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys	502
Db	1495	TTCCAGTTTACGTGCAAGCACATGAGGATGAGGACGACCTCAAGAGGGGGGAAG	1554
Qy	503	LysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSer	522
Db	1555	AAGCCCTTTGGGCACATTTGCACAGCAGCCACCCCTCTGCTGTGCACCGTCTCCAACAGC	1614
Qy	523	SerTrpAsnCysAspGlyValLeuHisSerProAlaIleGluValArgValHisCys	542
Db	1615	TCCTGGAATCGGATGGGAGGTCTGCACAGCCCTGCCATCGAGGTCAAGTCCACTGC	1674
Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer	562
Db	1675	CAGCTGGTTCGACTCTTTGCACGAGGAATTGAAGAGATCCGAAGCCAGACTCACACAGC	1734
RESULT 4			
LOCUS	BD102675	4463 bp	linear
DEFINITION	Ceramide kinase and DNA thereof.		
ACCESSION	BD102675		
VERSION	BD102675.1	GI:22648249	
KEYWORDS	WO 0196575-A/1.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Suglira, M., Kono, K. and Kohama, T.		
TITLE	Ceramide kinase and DNA thereof		
JOURNAL	Patent: WO 0196575-A 1 20-DEC-2001;		
COMMENT	SANKYO CO LTD, MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA		
	OS	Homo sapiens	(human)
	PN	WO 0196575-A/1	
	PD	20-DEC-2001	
	PF	11-JUN-2001	WO 2001JP004889
	PI	14-JUN-2000	JP OOP 178039
	PR	MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA	
	PC	C12N15/54, C12N9/12, C12N1/21, C12Q1/48, C07K16/40, A61K31/7125, PC A61K48/00,	
	PC	A61P3/04, A61P3/10, A61P7/00, A61P25/00, A61P29/00, A61P31/18, PC A61P35/00	
	CC	Ceramide kinase and DNA thereof	
	FH	Key	Location/Qualifiers
	FT	CDS	(124)..(1734).
FEATURES	Location/Qualifiers		

source	1. .4463	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Alignment Scores:			
Pred. No.:	9.55e-198	Length:	4463
Score:	3004.00	Matches:	558
Percent Similarity:	99.82%	Conservative:	1
Best Local Similarity:	99.64%	Mismatches:	1
Query Match:	99.31%	Indels:	0
DB:	6	Gaps:	0
US-10-631-958-11 (1-562) x BD102675 (1-4463)			
Qy	3	AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerPro	22
Db	55	GCCGCTAACGTCGCGGGCCCTCGGGCTCGCGGGCCCCCAGGCTGGCGGACGAGCCCG	114
Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys	42
Db	115	GCAGCGGAGATGGGGCGACGCGGGCGCGGAGCCGCTGCAATCCGTCTGTGGGTGAAG	174
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer	62
Db	175	CAGCAGCGCTCGCGCGTGGAGCTGGAGCCCGCGGGGCTCTGCTGCGCTGGTGGCGGAGC	234
Qy	63	ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle	82
Db	235	CCGGGGCCGAGCGCGGGCCCCCGGGCGCGATGCTGCTCTGTGCTGTATCTGAGATC	294
Qy	83	IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpLys	102
Db	295	ATCCGCGTTGAGGAAAACAGAGCTTTCACGGGAAAACATCAAGGAGTGGGAAAATGGCAGAA	354
Qy	103	MetGluLysProTyraAlaPheThrValHisCysValLysArgAlaArgHisArgTrp	122
Db	355	ATGAAAAGCCTTACGCTTTTACAGTTTCTGTTAAAGAGACGACGACGCGCTGG	414
Qy	123	LysTrpAlaGlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGln	142
Db	415	AAGTGGCGCAGGTGACTTCTGCTGTCAGAGAGGAGCAGTGTGTCTACTTGTGGCTGAG	474
Qy	143	ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle	162
Db	475	ACCTGCGGGAGATGCTGGAGAACTGACGCTCCAGACCAAGCATTTACTGGTATTTATC	534
Qy	163	AsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrgluArgLysValAlaProLeu	182
Db	535	AACCCGTTTGGAGGAAAAGGCAAGGCAAGCGGATATATGAAGAAAAGTGGCACCTG	594
Qy	183	PheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLys	202
Db	595	TTCACTTTAGCCTCCATCACCATGACATCATCTGTTACTGAACATGCTTAATCAGGCCAAG	654
Qy	203	GluThrLeuTyrgluIleAsnIleAspLysTyraAspGlyIleValCysValGlyGlyAsp	222
Db	655	GAGACTCTGATGAGATTAAACATAGACAAATACACGCGCATCGTCTGTGCGGGAGAT	714
Qy	223	GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal	242
Db	715	GGTATGTTACGAGAGTGTGTCACGGTGTGATTGGGAGGACGAGAGAGCGCGGGGTC	774
Qy	243	AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro	262
Db	775	GACCAGAACACCCCGGGCTGTGCTGCTCCCATAGTAGCTTCCGGATTGGAAATCATTCCC	834
Qy	263	AlaGlySerThrAspCysValCysTyraSerThrValGlyThrSerAspAlaGluThrSer	282
Db	835	GCAGGTCCACGGACTGGTGTGTGTACTCCACCGTGGGACACGACGACGAGAACCTCG	894
Qy	283	AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn	302

Db	895	GCCTGCATATCGTTGTTGGGACTCGCTGCCCATGATGTCTCCTCAGTCCACCAAC	954	
Qy	303	SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle	322	
Db	955	AGCACATCTCTTCGCTACTCCGTCCTCGTGGCTACGGCTTCTACGGGGACATCATC	1014	
Qy	323	LysAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys	342	
Db	1015	AAGACACAGTGAAGAAGACGGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAG	1074	
Qy	343	ThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr	362	
Db	1075	ACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACG	1134	
Qy	363	ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer	382	
Db	1135	GTGGGATCTCCAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGAGGCAAGC	1194	
Qy	383	LysGlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp	402	
Db	1195	AAGCAGCAGCTGGAGGAGGACAGAAAGACACTGTATGGTTTGGAGCTGCGGAGAC	1254	
Qy	403	ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer	422	
Db	1255	GTGGAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC	1314	
Qy	423	CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer	442	
Db	1315	TGTGCTTGTCCCGGAGCCCGAGGGGCTCTTCCCGGGCTGCCACTTGGGAGACGGGTCT	1374	
Qy	443	SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg	462	
Db	1375	TCTGACCTCATCTCATCCGGAATGCTCCAAAGTTCAATTTTCTGAGATTTCTCATCAGG	1434	
Qy	463	HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys	482	
Db	1435	CACACCAACGACGAGGACAGTTTGTACTTCTACTTTTGTGAAGTTTATCGCGTCAAGAA	1494	
Qy	483	PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys	502	
Db	1495	TTCCAGTTTACGTCGAAGCACATGAGGATGAGGACAGCGACCTCAAGAGGGGGGGAAG	1554	
Qy	503	LysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSer	522	
Db	1555	AAGGCTTTGGGCACATTTGGCAGGACGACCCCTCTCTGCTGCTGCACCGTCTCCAAACAGC	1614	
Qy	523	SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys	542	
Db	1615	TCCTGGAACTCGGATGGGAGGTCTTGCCACAGCCCTGCATCGAGGTCCAGAGTCCACTGC	1674	
Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluGlnAsnProLysProAspSerHisSer	562	
Db	1675	CAGCTGGTTGACTCTTTGACAGGGAATTGAAGAGATCCGAAAGCCAGACTCACACAGC	1734	
RESULT 5				
AR541900				
LOCUS	AR541900	4432 bp	DNA	linear
DEFINITION	Sequence 148 from patent US 6743619.			
ACCESSION	AR541900			
VERSION	AR541900.1	GI:539333980		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 4432)			
AUTHORS	Tang, Y. T., Zhou, P., Goodrich, R., Liu, C., Asundi, V., Ren, F., Zhang, J., Zhao, O. A., Yang, Y., Xue, A. J., Wehrman, T., Wang, J. R., Wang, D. and Drmanac, R. T.			
TITLE	Nucleic acids and polypeptides			
JOURNAL	Patent: US 6743619-A 148 01-JUN-2004;			
FEATURES	Location/Qualifiers			
	1..4432			
source				

/organism="unknown"					
/mol_type="genomic DNA"					
ORIGIN					
Alignment Scores:					
Pred. No.:	9.66e-193	Length:	4432		
Score:	2931.50	Matches:	548		
Percent Similarity:	98.21%	Conservative:	2		
Best Local Similarity:	97.86%	Mismatches:	7		
Query Match:	96.91%	Indels:	3		
DB:	6	Gaps:	1		
US-10-631-958-11 (1-562) x AR541900 (1-4432)					
Qy	6	GlyProAlaProLeuGly-----ValArgAlaProProAlaTrpArgThrSerPro	22		
Db	1135	GGTTCGCCCTATAGACAAGACAGTCAAGAGAGCTGGCAGGCTCGCGAGCAGGCCCG	1194		
Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys	42		
Db	1195	GCGGCGGAGATGGGGCGACGCGGGCGGCGGAGCCGCTGCAATCCGTGCTGTGGGTGAAG	1254		
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer	62		
Db	1255	CAGCAGCGCTCGCGCTGAGCTGGAGCCGCGGGCTCTGCTGCGCTGTGTGGCGGAGC	1314		
Qy	63	ProGlyProGlyAlaGlyAlaProGlyValAspAlaCysSerValProValSerGluIle	82		
Db	1315	CCGGGGCCGGAGCCGGCGGCCCCCGGGCGGATGCTCTGTGCTGTATCTGAGATC	1374		
Qy	83	IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys	102		
Db	1375	ATCGCGCTTGAGGAAACAGACGTTTACGCGGAAACATCAAGGCAGTGGAAAAATGGCAGAAA	1434		
Qy	103	MetGlnLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrp	122		
Db	1435	ATGGAAGGCTTACGCTTTTACAGTTTCTGTTAAAGAGAGCAGCAGCGCACCCCTGG	1494		
Qy	123	LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGln	142		
Db	1495	AAGTGGGCGCAGGTGACTTCTGCTGTCTCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAG	1554		
Qy	143	ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle	162		
Db	1555	ACCTCGCGGAGATGCTGGAGAGCTCAGCTCCAGACCAAGCAATTTACTTGGTATTATTATC	1614		
Qy	163	AsnProPheGlyLysGlyGlnGlyArgIleTyrGluArgLysValAlaProLeu	182		
Db	1615	AACCCGTTTGGAGGAAAGGACACAGGCAAGCGGATATATGAAAGAAAAGTGGCACCACTG	1674		
Qy	183	PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys	202		
Db	1675	TTCACTTAGCTCCATCACCATGACATCATCTGTTACTGAAACATGCTAATCAGGCCAAG	1734		
Qy	203	GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp	222		
Db	1735	GAGACTCTGTATGAGATTAAACATAGACAAATACACGGCATCGTCTGTCTGGCGGAGAT	1794		
Qy	223	GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal	242		
Db	1795	GGTATGTTTCAGCGAGGTGCTGCAAGGCTCTGATTGGGAGGACGACAGAGAGCGCGGGGTC	1854		
Qy	243	AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro	262		
Db	1855	GACCAGAACCAACCCCGGGCTGTGCTGGTCCCCAGTAGCTCCGGATTGGNATCATTTCCC	1914		
Qy	263	AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer	282		
Db	1915	GCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACACAGCGACGAGAAACCTCG	1974		
Qy	283	AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn	302		
Db	1975	GCGCTGCATATCGTTGTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAAC	2034		

Homo sapiens cDNA sequence. This sequence was generated as part of the Wellcome Trust Sanger Institute program to isolate cDNA clones representing the full length open reading frame of well annotated protein coding genes on human chromosome 22. For more information see <http://www.sanger.ac.uk/HGP/Chr22/>.

FEATURES	SOURCE
1. Age	1. Age
2. Gender	2. Gender
3. Marital Status	3. Marital Status
4. Education	4. Education
5. Income	5. Income
6. Occupation	6. Occupation
7. Religion	7. Religion
8. Political Affiliation	8. Political Affiliation
9. Health Status	9. Health Status
10. Travel History	10. Travel History
11. Employment Status	11. Employment Status
12. Family Size	12. Family Size
13. Home Ownership	13. Home Ownership
14. Vehicle Ownership	14. Vehicle Ownership
15. Insurance Status	15. Insurance Status
16. Charitable Donations	16. Charitable Donations
17. Volunteer Work	17. Volunteer Work
18. Substance Use	18. Substance Use
19. Stress Levels	19. Stress Levels
20. Life Satisfaction	20. Life Satisfaction

```

1. .1654
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /chromosomes="22"
  /clone="pGEM.bk29F11.1"
  /lab_host="JM109"
1. .1654
  /gene="bk29F11.1"
11. .1624
  /gene="bk29F11.1"
  /codon_start=1
  /protein_id="CA3630290.1"
  /db_xref="GI:47678339"

```

ss. "MGATGAAEPTLSVLVWVKOORCAVLEPARALLRWRRSPGPGAGA
/translation="MGATGAAEPTLSVLVWVKOORCAVLEPARALLRWRRSPGPGAGA
PGDACSVPVSIATEEDVHGKSGSGKGMQKBPFTVHCVRKRRHRRWKAQV
TWCPCEQLCHLWTLREMLKLSRQKHLAVTFNFGQGSGLIYERKVAFLPT
ASITDITVTHANQAKETLVINIDKIDGIVCGVDGMSVHGLIGLQRTORSAGV
QNIIPRAVLPPSLRIGIIPAGSDICVCTVCGDGAETSAHLIYGLSDIAMDVSVVH
NSTLLYSVSLIGVGFYDILKDSKKQBLGLYARDPSGLTFLSHCYEGTVSLFLPA
QTVGSPDRDRCPKAGCFVCKSQKQLEGGKALYGLBAEDBHWVVCVKGKFLPA
ATNNSACDRSPGLSPAAHGLDGSDDLILTKCSRFPNRLRIHQNTODQDFTFV
EAYRVKFFQTSYKIMEDSDLKKEGGKRGFGHICSHSPSCCCTVSNSSNNSCDGEVLHS
PAIVRYHCOLVLPARGIEENLKKPKDHS"

ORIGIN

Alignment Scores:		
Pred. No.:	3.59e-191	Length: 1654
Score:	2901.00	Matches: 540
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	95.90%	Indels: 0
DB:	9	Gaps: 0

US-10-631-958-11 (1-562) x CR456404 (1-1654)

Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValIlys	42
Db	2	GCSCGCGAGATGCGGGCGACGCGGGCGCGAGCGCTGCAATCGTGCTGTGGGTGAAG	61
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer	62
Db	62	CAGCAGCGTGGCGGTGAGCTCGAGCCCGCGCGGGCTCTGCTGGCTGGTGGCGGAGC	121
Qy	63	ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle	82
Db	122	CCGGGCGCGAGCGCGCGCCCGCGCGCGATGCTCTCTGTGGCTGTATCTGAGATC	181
Qy	83	IleAlaValGluGluThrAspValHisGlyIleHisGlnGlySerGlyIleTrpGlnIys	102
Db	182	ATCGCGGTGGAGAAACAGAGCTTCACGGGAAACATCAAGGCAGTGGAAATGCGAGAA	241
Qy	103	MetGluIysProTyrAlaPheThrValHisCysValIysArgAlaArgHisArgTrp	122
Db	242	ATGGAAAGCCTTACGCTTTTACAGTTCTCACTGTGTAAAGAGAGCAGCACCGCTGG	301
Qy	123	LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGln	142
Db	302	AAGTGGCGCAGGTGACTTCTTGGTGTCCAGAGAGCAGCTGTGTCACTTGTGGCTCCAG	361
Qy	143	ThrLeuArgGluMetLeuGluIleLeuThrSerArgProLysHisLeuLeuValPheIle	162
Db	362	ACCTCGGGGAGATGCTGGAGAGCTGACGTCCAGCCAAAGCAATTTACTGGTATTATC	421
Qy	163	AsnProPheGlyGlyIleGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu	182
Db	422	AACCCGTTTGGAGGAAAGGCAAGCAGCGGATATATGAAAGAAAGTGGACCACTG	481

Qy	303	SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLe	322
Db	2035	AGCACATCTCTTCGCTACTCGTGTCCCTGTCTGGGCTACGGCTTCTACGGGGAATCATC	2094
Qy	323	LysAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys	342
Db	2095	AAGARCAAGTGAAGAANAACGGTGGTCTGCCAGATACGACTTTTCAGGTTTAAAG	2154
Qy	343	ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr	362
Db	2155	ACCTTCCTCTCCACCACTGCTATGGAGGGACAGTGTCTTCTCTCCCTGCACAAACACACG	2214
Qy	363	ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer	382
Db	2215	GTGGGATCTCCAAAGGGATAGGAACCCCTCCCGGCAGGATGCTTTGTGTGCAGGCAAAAGC	2274
Qy	383	LysGlnGlnLeuLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp	402
Db	2275	AAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAGGAC	2334
Qy	403	ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer	422
Db	2335	GTGGAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTC	2394
Qy	423	CysAlaCysArgSerProArgLysSerProAlaAlaHisLeuGlyAspGlySer	442
Db	2395	TGTGCTTGTGCGCCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGACGGGTCT	2454
Qy	443	SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg	462
Db	2455	TCTGACCTCATCTTCATCCGGAAATGCTCCAGGTTCAATTTCTCGAGATTTCTCATCAGG	2514
Qy	463	HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys	482
Db	2515	CACACCAACGACGACGACCAGTTTGACTTTCACCTTTGTTGAAGTTTATCGCTCAAGAA	2574
Qy	483	PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys	502
Db	2575	TTCCAGTTTACGTCCGAAGCACATGGAGGATGAGACACGACCTTCAAGGAGGGGGGGAAG	2634
Qy	503	LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSer	522
Db	2635	AAGCGCTTTGGGCACATTTGAGCAGCACCCCTCTCTGCTGCTGCACCGTCTCCAAACAGC	2694
Qy	523	SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys	542
Db	2695	TCCTGGAACTGCGACGGGGAGGTCTTGCAACAGCCCTGCCATCGAGGTCCAGTCCATGCG	2754
Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer	562
Db	2755	CAGCTGTTTGACTTCTTGCACGAGGAATTAAGAGAATTCGAAGCCGAGCTCACACAGC	2814

RESULT 6	
CR456404	1654 bp mRNA linear PRI 25-MAY-2004
LOCUS	Homo sapiens bk29F11.1 full length open reading frame (ORF) CDNA
DEFINITION	clone (cDNA clone C22ORF.pGEM.bk29F11.1).
ACCESSION	CR456404
VERSION	CR456404.1 GI:47678338
KEYWORDS	CDNA; chromosome 22; ORF.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1654) Collins,J.E., Wright,C.L., Edwards,C.A., Davis,M.P., Grinham,J.A., Cole,C.G., Goward,M.E., Aguado,B., Mallia,M., Mokrab,Y., Huckle,E.J., Beare,D.M. and Dunham,I.
TITLE	Direct Submission
JOURNAL	Submitted (24-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: c22q@sanger.ac.uk Manuscript
COMMENT	Sanger Institute name : pGEM.bk29F11.1

Qy	183	PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAenGlnAlaLys	202
Db	482	TTCACTTAGCTCCATCACCACCTGACATCATCGTTACTGAACATGCTAATCAGGCCAAG	541
Qy	203	GluThrLeuTyrGluIleAenIleAspLysTyrAspGlyIleValCysValGlyGlyAsp	222
Db	542	GAGACTGTATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTCTGGCGGAGAT	601
Qy	223	GlyMetPheSerGluValLeuHisGlyLeuIleGlyValArgThrGlnArgSerAlaGlyVal	242
Db	602	GGTATGTTTCAGCAGGTTGCTGCAGCGTCTGATTGGAGGACGACAGAGCGCGGGGTC	661
Qy	243	AspGlnAenHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro	262
Db	662	GACCAGAACCCACCCCGGGTGTGCTGCTCCAGTAGCCTCGGATTGGAATCAATCCC	721
Qy	263	AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer	282
Db	722	GCAGGGTCAACGGACTCGTGTGTACTCCACCGTGGGCACACGACGACGAGAACCTCG	781
Qy	283	AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn	302
Db	782	CGCTGCTATATCGTTGTGGGACTCGCTGGCCATGATGTGCTCAGTCCACCAAC	841
Qy	303	SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle	322
Db	842	AGCACACTCTTCGCTACTCCGTGCTCCTGCTGGGTACGGCTTCTACGGGGACATCATC	901
Qy	323	LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys	342
Db	902	AAGGACAGTGAAGAAACAGTGTGGTCTTGCCACAGATACGACTTTTCAGGTTTAAAG	961
Qy	343	ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr	362
Db	962	ACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACG	1021
Qy	363	ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer	382
Db	1022	GTGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCCAAGC	1081
Qy	383	LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp	402
Db	1082	AAGCAGCAGCTGGAGGAGGACGAGAAGAACACTGTATGTTTGGAAAGCTGCGGAGAC	1141
Qy	403	ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAenAlaThrAsnMetSer	422
Db	1142	GTGAGGAGTGGCAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCAACAATGTCC	1201
Qy	423	CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer	442
Db	1202	TGTGCTTGTCCCGAGCCCAAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGTCT	1261
Qy	443	SerAspLeuIleLeuIleArgLysCysSerArgPheAenPheLeuArgPheLeuIleArg	462
Db	1262	TCTGACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG	1321
Qy	463	HisThrAenGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys	482
Db	1322	CACACCAACGACGAGGACCACTTTCATCTTCTTGTGAAGTTTATCGCGTCAAGAAA	1381
Qy	483	PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys	502
Db	1382	TTCCAGTTTACGTCGAGACATGAGGATGAGGACAGCGACCTCAAGGAGGGGGGAG	1441
Qy	503	LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSer	522
Db	1442	AAGCGCTTGGGACATTTGCAGCAGCACCCCTCTCTGCTGCTCACCGTCTCCACAGC	1501
Qy	523	SerTrpAenCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys	542
Db	1502	TCCTGGAACTCGCAGCGGGAGGTCTGCACAGCCCTCCATCGAGGTCCACTGCTGC	1561

Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluGluAenProLysProAspSerHisSer	562
Db	1562	CAGCTGGTTCCACTCTTTTCACGAGGAATTGAAGAGAATCCGAAGCCAGACTCACAGC	1621
RESULT 7			
BC067255		1772 bp	linear
LOCUS		1772 bp	linear
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			

gene
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6185601"
/tissue_type="Peripheral Nervous System, dorsal root ganglion"
/clone_lib="Lupski_dorsal_root_ganglion"
/lab_host="DH108"
/note="Vector: PCMV-SPORT6"
1..1772
/gene="CERK"
/note="synonyms: KIAA1646, hCERK, dA59H18.3, DKFP434E0211, FLJ21430, FLJ23239, dA59H18.2, LK4"
/db_xref="LocusID:64781"
579..11598
/gene="CERK"
/codon_start=1
/product="CERK protein"
/protein_id="AAH67255.1"
/db_xref="GI:45595583"
/db_xref="LocusID:64781"
/translation="MFSVLHGLIGRTQRSAGVDNHPRAVLVPSSLRIGIIPAGSTD
CVSVSTGSDAETSAIHLVVGDSLMDVSSVHHNSTLLRYSVSLGYFGYDIIKDS
EKKRLWGLADFSLGKTLFSLHCVGVTSFLPAQHTVGSPRDRKPCRCAGFCVCRQSK
OOLEEOKALYGLBAADVEEMOVVCGKFLAINATNMSCACRRSPRGLSPAHLGDG
SGDLILPKSRFENFLRLIHTNQDQDFTEFEVYVKFKFQFTSKHWEDESDLKE
GKKRPFHICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIENPK
PDSHS"

CDS
Alignment Scores:
Pred. No.: 1,238-183 Length: 1772
Score: 2792.50 Matches: 530
Percent Similarity: 95.15% Conservative: 0
Best Local Similarity: 95.15% Mismatches: 1
Query Match: 92.31% Indels: 27
DB: 9 Gaps: 2

ORIGIN

US-10-631-958-11 (1-562) x BC067255 (1-1772)
QY 7 ProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerProAlaAlaGluMet 26
DB 3 CGCGCGCCCTCGCGGTCCGCGCGCCCGCCCGCGCGCGCGCGCGCGCGCGCGCGGATG 62
QY 27 GlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValGlnGlnArgCys 46
DB 63 GCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122
QY 47 AlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyProGly 66
DB 123 GCGGTGAGCTGGAGCCCGCGCGCGCTGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGGA 182
QY 67 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaValGlu 86
DB 183 GCGCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 242
QY 87 GluThrAspValHisGlyHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro 106
DB 243 GAAACAGACGTTTCACGGGAAACATCAAGCAGTGGGAAATGGCAGAAATGGAAAGCCT 302
QY 107 TyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAlaGln 126
DB 303 TACGCTTTTACAGTTTACCTGTGTAAAGAGAGCAGCGCGCACCGCTGGAAGTGGCGGAG 362
QY 127 ValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu 146
DB 363 GTGACTTCTGTGTCCAGAGAGCAGCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 422
QY 147 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 166
DB 423 ATGCTGGAGAAAGCTGACGCTCCAGACCAAGCAATTTTACTGGTATTATCAAC ----- 473
QY 167 GlyLys-GlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAl 186

DB 474 ---AAAAGGACAAGCGCGGATATATGAAGAAAGAGTGGCACCACCTGTTTCACTTACG 530
QY 186 aSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTy 206
DB 531 CTCCATCACCCTGACATCATCGC----- 554
QY 206 rGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSe 226
DB 555 -----ATCGTCTGTGTCGGCGGAGATGGTATGTTTCAAG 586
QY 226 rGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHi 246
DB 587 CGAGGTCTCAGCGGTCTGATTGGAGGACGACGAGGAGCGCGGGGTTCGACCAACCA 646
QY 246 sProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerTh 266
DB 647 CCCCCTGGCTGTGCTGGTCCCGTAGCTCCGGATTGGAATCATTCCTCCGACGGGTCAAC 706
QY 266 rAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisI 286
DB 707 GGACTGCGTGTGTACTCCACCGTGGGCACGAGGACGACGAAACCTCGCGCTGCATAT 766
QY 286 eValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSerThrLeuLe 306
DB 767 CGTTGTTGGGACCTCGCTGGCCATGGATGTCTCTCAGTCCACCAACACGACACTCTCT 826
QY 306 uArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGl 326
DB 827 TCGCTACTCCGTGCTCCCTGCTGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGA 886
QY 326 uLysIleArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSe 346
DB 887 GAAGAAACGGTGGTGTGGCTTGTCCAGATACGACTTTTTCAGGTTTAAAGAGCTTCTCTC 946
QY 346 rHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerPr 366
DB 947 CCACCACTGCTATGAGGACAGTGTCTCTCTCTCTGTCACACACACGCTGGGATCTCC 1006
QY 366 oArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLe 386
DB 1007 AAGGGATAGGAAGCCCTGCGCGGCGAGGATGCTTTGTTGACGGAAGCAAGCAGCAGCT 1066
QY 386 uGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluTr 406
DB 1067 GGAGGAGGAGCAGAGAAAGCACTGTATGGTTCGGAAGCTCGCGAGGACGTGGAGGAGTG 1126
QY 406 pGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysAr 426
DB 1127 CCAAGTCTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTGTCTGTCTG 1186
QY 426 gArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuI 446
DB 1187 CCGGAGCCCCAGGGGCTCTCCCGGCTGCCCCACTTGGGAGACGGGTCTTCTGACCTCAT 1246
QY 446 eLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGl 466
DB 1247 CCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGSCACCAACCA 1306
QY 466 nGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPheTh 486
DB 1307 CGAGACCAAGTTTGAATTTTGTGTAAGTTTATCGCGTCAAGAAATTTCCAGTTTAC 1366
QY 486 rSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPheGl 506
DB 1367 GTCCAGACACATGAGGATGAGGACGACGCTCAAGGAGGGGGGGAAGAGCGCTTTGG 1426
QY 506 yHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCy 526
DB 1427 GCACATTTGACAGCAGCCCTCTCTGCTGTGTCACGCTCTCCAAACAGCTCTCTCGAACTG 1486
QY 526 sAspGlyValValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValAr 546
DB 1487 CGAGGGGAGGTCTCTGACAGCCCTGCGCATCGAGTCAGATTCACCTCCAGCTGGTTGG 1546

```
Qy 546 gLeuPheAlaArgGlyIleGluGluAenProLysProLysProAspSerHisSer 562
Db 1547 ACTCTTTGCACGAGGAATTGAAGAGAAATCCGAAGCCAGACTCACACAGC 1595

RESULT 8
AX224383 1840 bp DNA linear PAT 10-SEP-2001
LOCUS Sequence 5 from Patent WO0160990.
DEFINITION AX224383
ACCESSION AX224383
VERSION AX224383.1 GI:15554633
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Rastelli, L.
TITLE Novel sphingosine kinases
JOURNAL Patent: WO 0160990-A 5 23-AUG-2001;
Curagen Corporation (US); GENENTECH, INC. (US)
FEATURES
source
1..1840
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 9.77e-174 Length: 1840
Score: 2649.50 Matches: 495
Percent Similarity: 97.83% Conservative: 1
Best Local Similarity: 97.63% Mismatches: 0
Query Match: 87.59% Indels: 11
DB: 6 Gaps: 1

US-10-631-958-11 (1-562) x AX224383 (1-1840)

Qy 67 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaValGlu 86
Db 1 GCCGGCGCCCGCGGGGATGCTCTGTGCTGTATCTGAGATCATCGCGTTGAG 60
Qy 87 GluThrAspValHisGlyHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro 106
Db 61 GAAACAGACGTTACCGGGAACATCAAGCGAGTGGAAAATGGCAGAAAATGGAAAAGCCT 120
Qy 107 TyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAlaGln 126
Db 121 TACGCTTTTACAGTTCACTGTGTAAGAGAGACGACGCGCACCCGCTGGAAAGTGGCGCAG 180
Qy 127 ValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu 146
Db 181 GTGACTTTCTGGTGTCCAGAGGACAGCTGTGTCTACCTTGTGGCTGCAGACCCCTCGGGAG 240
Qy 147 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 166
Db 241 ATGCTGAGAGAGCTGACGTCACAGCAACCAAGCATTTACTGGTATTATCAACCCCGTTTGA 300
Qy 167 GlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAla 186
Db 301 GGAAGAAGGACAAGCAAGCGGATATATGAAAGAAAAGTGGCACCACCTGTTCACCTTAGCC 360
Qy 187 SerIleThrThrAspIleIle-----ValThr 195
Db 361 TCCATCACCCTGACATCATCGGTAAACAAATTCATGTGTAACATGTAGAGTAATTACT 420
Qy 196 GluHisAlaAenGlnAlaLysGluThrLeuTyrGluIleAenIleAspLysTyrAspGly 215
Db 421 GAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGCGC 480
Qy 216 IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyIleGlyArg 235
Db 481 ATCGTCTGTGCGCGGAGATGGTATGTTTCAGCAGGCTGTCACCGGCTGATTGGGAGG 540

236 ThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSer 255
541 ACCGAGAGGAGCGCGGGTCGACCAGAACACCCCGGGCTGTGTGGTCCCCAGTAGC 600
256 LeuArgIleGlyIlelleProAlaGlySerThrAspCysValCysTyrSerThrValGly 275
601 CTCGGATTGGAATCATTCGCGAGGGTCAACGACTGCGTGTGTACTCCACCGTGGGC 660
276 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAsp 295
661 ACCAGCGACGCAAAACCTCGGCGCTGCATATCGTTTGGGACTCGCTGGCCATGGAT 720
296 ValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyr 315
721 GTGCTCTCAGTCCACCACACAGACACATCTCTTCGCTACTCCGTGTCTCTGGGCTAC 780
316 GlyPheTyrGlyAspIlelleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArg 335
781 GGGTCTTACCGGGACATCATCAAGGACAGTGAAGAAACGTTGGTGGTCTTGGCCAGA 840
336 TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSer 355
841 TAGACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCC 900
356 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 375
901 TTCTCTCTCGCACACACACGCTGGGATCTCCAAGGATAGGAAGCCCTCCCGGCAGGA 960
376 CysPheValCysArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyr 395
961 TGCTTTGTTTGCAGGCAAAACAAAGCAGCAGCTGGAGAGAGAGAGAGAAAGCACTGTAT 1020
396 GlyLeuGluAlaAlaGluAspValGluTrpGlnValValCysGlyLysPheLeuAla 415
1021 GGTTTGAAGCTCGGAGGACGTGGAGAGTGGCAAGTCTGTGTGGAAATTTCTGGCC 1080
416 IleAsnAlaThrAsnMetSerCysAlaCysArgSerProArgGlyLeuSerProAla 435
1081 ATCAATGCCACAAACATGCTCTGTGCTGTTCGCGGAGCCCGAGGGCCCTCTCCCGGCT 1140
436 AlaHisLeuGlyAspGlySerSerAspLeuIleLeuLysLysCysSerArgPheAsn 455
1141 GCCCACTTGGGAGACGGTCTTCTGACCTCATCTCATCCGAAATGCTCCAGGTTCAAT 1200
456 PheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheVal 475
1201 TTTCTGAGATTCTCATCAGCACACCAACACAGGACCCAGTTTGATCTTCACTTTTGT 1260
476 GluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSer 495
1261 GAAGTTTATCGCTCAAGAAATTCAGTTTACGTTCGAAGCACATGGAGGATGAGGACAGC 1320
496 AspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCys 515
1321 GACCTCAAGAGAGGGGGAAGAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCTCGC 1380
516 CysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAla 535
1381 TGTGTCACCGCTCTCCAACAGCTCTTGAATTCGACGGGAGGCTCTGCACAGCCCTGCC 1440
536 IleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsn 555
1441 ATCGAGGTGAGGTCCACTGCCAGTGTTCGACTCTTTTCACGAGGAATTTGAAGAGAT 1500
556 ProLysProAspSerHisSer 562
1501 CCGAAGCCAGACTCACACAGC 1521

RESULT 9
BD183468
LOCUS BD183468
DEFINITION Novel genes and proteins encoded by the genes.
PAT 17-JUN-2003
```

```

ACCESSION   BD183468
VERSION     BD183468.1  GI:31875668
KEYWORDS    JP 2002345492-A/181.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 4171)
AUTHORS     Ohara, O., Nagase, T. and Nakajima, D.
TITLE       Novel genes and proteins encoded by the genes
JOURNAL     Patent: JP 2002345492-A 181 03-DEC-2002;
            KAZUSA DNA RESEARCH INSTITUTE
COMMENT     OS Homo sapiens (human)
            PN JP 2002345492-A/181
            PD 03-DEC-2002
            PF 26-FEB-2002 JP 2002049009
            PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
            PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61P25/00,
            A61P25/14,
            PC A61P25/18, A61P35/00, C12N15/00, A61K37/02
            CC Novel genes and proteins encoded by the genes FH Key
            Location/Qualifiers
            FT CDS
               (2)..(1444).
FEATURES             Location/Qualifiers
     source           1..4171
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:          4,848-169      Length:          4171
Score:              2588.00        Matches:         481
Percent Similarity: 100.00%        Conservative:    0
Best Local Similarity: 100.00%      Mismatches:     0
Query Match:        85.55%          Indels:          0
DB:                 6              Gaps:            0

US-10-631-958-11 (1-562) x BD183468 (1-4171)
QY      82  lleileAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGln 101
Db      2  ATCATCGCGTGTGAGGAAACAGACGTTTACGGGAAACATCAAGGAGTGGAAATGGCAG 61
QY      102 LysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArg 121
Db      62  AAAATGGAAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCACCGC 121
QY      122 TrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuLeuCysHisLeuTrpLeu 141
Db      122 TGGAAAGTGGCGCGAGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTTGTGGCTG 181
QY      142 GlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPhe 161
Db      182  CAGACCCCTGCGGGAGATGCTGGAAGAGCTGAGCTCCAGACCAACATTTACTGTATTT 241
QY      162 lleAsnProPheGlyGlyGlnGlyLysArgGlieTyrGluArgLysValAlaPro 181
Db      242  ATCAACCCGTTTGGAGGAAAGGACAGCGGATATATGAAAGAAAGTGGCACCA 301
QY      182 LeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAla 201
Db      302  CTGTTTCACTTAGCCCTCCATCACCACCTGACATCATCTGTTACTTGAACATGCTAATCAGGCC 361
QY      202 LysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGly 221
Db      362  AAGGAGACTCTGTATGAGATTAACTAGACAATACGCGGATCGTCTGTGTGCGCGGA 421
QY      222 AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly 241
Db      422  GATGTATGTTTCAGCGAGGTGTCGACGCTCTGATTTGGAGGACGACGAGGAGCGCCGGG 481
QY      242 ValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 261

```

```

Db      482  GTCGACCAAGAACACCCCGGGCTGTGCTGTGCCAGTAGCCTCCGATTTGGAAATCATTT 541
QY      262  ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr 281
Db      542  CCCGACGGGTCAACGGACTGCGTGTGTACTCCACCGTGGGCACCGACGACGACGACAAACC 601
QY      282  SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHis 301
Db      602  TCGCGCTGCATATCGTTTGGGACTCGCTGCCCATGGATGTGTCTCAGTCCACCAC 661
QY      302  AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 321
Db      662  AACAGCACATCTTCCGTACTCCGTCTCTGGCTGTGGCTACGGCTTCTACGGGACATC 721
QY      322  IleLysAspSerGluLysLysLeuArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 341
Db      722  ATCAAGGACACGTAGAGAAACGGTGGTGGTCTTGGCAGATACGACTTTTCAGGTTTA 781
QY      342  LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 361
Db      782  AAGACCTTCTCTCCACCACTGCTATGAAGGGACAGTGTCTCTCTCCCTGCACAACAC 841
QY      362  ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 381
Db      842  ACGTGGGATCTCCAAGGATAGGAGCCCTGCCGGCAGGATGCTTTTGTTCAGGCAA 901
QY      382  SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu 401
Db      902  AGCAAGCAGCAGCTGGAGGAGCAGAGAAGAACACTGTATGTTTGAAGCTGCGGAG 961
QY      402  AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet 421
Db      962  GACGTGGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCCAACAAATG 1021
QY      422  SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 441
Db      1022  TCCTGTGCTGTGCGCGAGGCCCTCTCCCGGCTGCCCATTTGGAGACGGG 1081
QY      442  SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle 461
Db      1082  TCTTCTGACCTCATCTCTCATCGGAAATGCTCCAGGTTCAATTTTCTCAGATTTCTCATC 1141
QY      462  ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgVallys 481
Db      1142  AGGCACACCAACACGACGACGACAGTTCAGTTCATCTTTTGTGAAGTTTATCGCGTCAAG 1201
QY      482  LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGly 501
Db      1202  AAATTCAGTTTACGTGGAAGCACATGGAGATGAGGACGACGACCTCAAGAGGGGGGG 1261
QY      502  LysLysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsn 521
Db      1262  AAGAAAGCGCTTTGGGCACATTTTGACGACGACACCCCTCTCTGCTGCTGACGCTCCAAAC 1321
QY      522  SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHis 541
Db      1322  AGCTCCTGGAACCTCGACGGGAGGTCCTGCACAGCCCTGCGCATCGAGGTCCAGAGTCCAC 1381
QY      542  CysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHis 561
Db      1382  TGCAGCTGTGTTCGACTCTTTGTCACGAGGAATGAAGAGAATCCGAGCCAGACTCACAC 1441
QY      562  Ser 562
Db      1442  AGC 1444

```

```

RESULT 10
AB051433
LOCUS    AB051433
DEFINITION Homo sapiens mRNA for KIAA1646 protein, partial cds.
ACCESSION AB051433
VERSION   AB051433.1  GI:13359166

```


KEYWORDS	1	162	IleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaPro	181
SOURCE	2	242	ATCAACCCGTTTGGAGAAAGGACAAAGCGCATATATGAAGAAAAGTGGCACCA	301
ORGANISM	3	182	LeuPheThrIleuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAla	201
REFERENCE	4	302	CTGTTACCTTAGCTCCATCACCCTGACATCATCTGTTACTGAACATGCTTAATCAGGCC	361
AUTHORS	5	202	LysGluThrIleuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGly	221
TITLE	6	362	AAGGAGACTCTGATGAGATTAAATAGACAAATACAGCGCATCTGCTCTGTGCGCGGA	421
JOURNAL	7	222	AspGlyMetPheSerGluValLeuHisGlyIleGlyArgThrGlnArgSerAlaGly	241
MEDLINE	8	422	GATGGTATGTTACGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGGG	481
PUBMED	9	242	ValAspGlnAsnHisProArgAlaValLeuValProSerSerIleuArgIleGlyIleIle	261
REFERENCE	10	482	GTCGACCAAGAACCCCGCGGCTGCTGCTGCCATGAGTACCTCCGATTTGAATCAT	541
AUTHORS	11	262	ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr	281
TITLE	12	542	CCGCGAGGTTCAACGAGTCTGCTGTTACTCCACCGTGGGACACGACGACGAGAAC	601
JOURNAL	13	282	SerAlaLeuHisIleValValGlyAspSerIleuAlaMetAspValSerSerValHisHis	301
MEDLINE	14	602	TCGGCGCTGCATATCGTTGTTGGGACTCGCTGCCATGATGTGCTCCTCAGTCCAC	661
PUBMED	15	302	AsnSerThrIleuLeuArgTyrSerValSerIleuGlyTyrGlyPheTyrGlyAspIle	321
REFERENCE	16	662	AACAGCACACTCTCTGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	721
AUTHORS	17	322	IleIleAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeu	341
TITLE	18	722	ATCAAGGACAGTGGAGAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	781
JOURNAL	19	342	LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis	361
MEDLINE	20	782	AAGACCTTCTCTCTCCACCACTGCTATGAAGGAGACAGTGTCTCTCTCTCTCTCT	841
PUBMED	21	362	ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln	381
REFERENCE	22	842	ACGGTGGGATCTCAAGGGATAGAAAGCCCTGCGCGGACAGTCTGTTGTTGCGAGCA	901
AUTHORS	23	382	SerLysGlnGlnLeuGluGluGlnLysAlaLeuLysAlaLeuTyrGlyLeuGluAlaGlu	401
TITLE	24	902	AGCAAGCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGGTTGGAGTGGCGAG	961
JOURNAL	25	402	AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet	421
MEDLINE	26	962	GACGTGGAGGAGTGGCAAGTCTGTTGGGAAGTTTCTGGCCATCAATGCCACAAATG	1021
PUBMED	27	422	SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly	441
REFERENCE	28	1022	TCCTGTGCTTGTGCGCGAGCCCGCGGCTCTCTCCCGGCTGCTGCTGCTGCTGCTG	1081
AUTHORS	29	442	SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle	461
TITLE	30	1082	TCCTGTGACCTCATCTCTATCCGAAATGCTCCAGTTCAATTTCTGAGATTTCTCATC	1141
JOURNAL	31	462	ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys	481
MEDLINE	32	1142	AGGCACACCAACACGACGAGGACCATTTGACTTCTACATTTTCTGAAAGTTTATCGCGT	1201
PUBMED	33	482	LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly	501
REFERENCE	34	1202	AAATTCAGTTTACGTCGAAGCACTGAGGAGATGAGACAGCGACCTCAAGAGGGGGGG	1261
AUTHORS	35	502	LysLysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsn	521
TITLE	36	1262	AAGAGCGCTTTGGGCACATTTGACGACGACACCCCTCTCTGCTGCTGCTGCTGCTG	1321
JOURNAL	37	522	SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHis	541

ORIGIN

Alignment Scores:
 Pred. No.: 4,84e-169 Length: 4171
 Score: 2588.00 Matches: 481
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.55% Indels: 0
 DB: 9 Gaps: 0

US-10-631-958-11 (1-562) x AB051433 (1-4171)

Qy	82	IleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGln	101
Db	2	ATCATCCCGTTGAGGAACACAGCTTTCACGGGAAACATCAAGGACGTGGAAAAATGGCAG	61
Qy	102	LysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArg	121
Db	62	AAATGGAAGACCTTACGCTTTTACAGTTTCACTGTGTGAAGAGGACGACGACCGC	121
Qy	122	TrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeu	141
Db	122	TGGAAGTGGGCGCAGGTGACTTTCTGCTGTCACAGGAGCAGCTGTCTCACTTGTGGCTG	181
Qy	142	GlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPhe	161
Db	182	CAGACCCCTGGCGGAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTTACTGGTATT	241


```

Db      1322  AGCTCTGGAAGTGGACGGGAGGTCTTGCACACCCCTGCCATCGAGGTACAGTCCAC 1381
Qy      542   CysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAAsnProLysProAspSerHis 561
Db      1382  TGGCAGCTGGTTCGACTCTTTTGCACGAGGAATTGAAGAGATCCGAGGCCAGACTCACAC 1441
Qy      562   Ser 562
Db      1442  AGC 1444

RESULT 11
CQ730476
LOCUS      CQ730476                1459 bp    DNA        linear    PAT 03-FEB-2004
DEFINITION Sequence 16410 from Patent WO02068579.
ACCESSION  CQ730476
VERSION    CQ730476.1  GI:42304409
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE     Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL   Patent: WO 02068579-A 16410 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES   Location/Qualifiers
            source
            1..1459
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.:      3,17e-164      Length:      1459
Score:          2510.00      Matches:      478
Percent Similarity: 87.89%      Conservative: 1
Best Local Similarity: 87.71%      Mismatches: 6
Query Match:      82.98%      Indels:      61
DB:              6          Gaps:      3

US-10-631-958-11 (1-562) x CQ730476 (1-1459)

Qy      18   TrpArgThrSerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSer 37
Db      3    TGGCGGACGAGCCCGCGCGGAGATGGGGCGACGGGGCGCGGAGCCCGCTGCATCC 62
Qy      38   ValLeuTrpValLysGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeu 57
Db      63   GTGCTGTGGTGAAGCAGCAGCGCTGCACAA----- 94
Qy      58   ArgTrpTrpArgSerProGlyAlaGlyAlaProGlyAlaAspAlaCysSerVal 77
Db      95   -----GGACGTGGTGT-----TTTACAGATGCCTGTCTGTG 127
Qy      78   ProValSerGluIleAlaValGluThrAspValHisGlyLysHisGlnGlySer 97
Db      128  CTGTATCTGAGATCATCGCGTTTGAAGAAACAGACGTTTACGGTTTACAGTTCACTGTGTAAGAGACGA 187
Qy      98   GlyLysTrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAla 117
Db      188  GGAAATGGCAGAAATGGAAAGCCTTACGCTTTTACAGTTCACTGTGTAAGAGACGA 247
Qy      118  ArgArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCys 137
Db      248  CGACGGCACCCTGGAAGTGGCGCGAGGTGACTTTCTGTGTGTCCAGAGGAGCAGCTGTGT 307
Qy      138  HisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHis 157
Db      308  CACTTGTGGCTGCAGACCCCTCGGGAGATGTGGAGAGCTG----- 349

```

```

Qy      158  LeuLeuValPheIleAAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArg 177
Db      349  ----- 349
Qy      178  LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHis 197
Db      350  -----ATTACTGAACAT 361
Qy      198  AlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleVal 217
Db      362  GCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTC 421
Qy      218  CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln 237
Db      422  TGTGTCCGCGGAGATGGTATGTTTCCAGCGAGGTGTCACGGTCTGATTGGGAGGACGCG 481
Qy      238  ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg 257
Db      482  AGGAGCGCGGGGTGCGACGACACACCCCGGGCTGTGCTGCTCCCGTAGCTCCGG 541
Qy      258  IleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSer 277
Db      542  ATTGGAATCATTTCCGCGAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGCACACG 601
Qy      278  AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSer 297
Db      602  GACGCAAAACCTCGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCC 661
Qy      298  SerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 317
Db      662  TCAGTCCACCACAAACAGCACACTCTTCGCTACTCCGCTGCTCCGCTGGGCTACGGCTTC 721
Qy      318  TyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAsp 337
Db      722  TACGGGACATCATCAGGACAGTGAAGAAACGGTGTGTTGGTCTTCCACAGATACGAC 781
Qy      338  PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu 357
Db      782  TTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTC 841
Qy      358  ProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 377
Db      842  CCTGCACAACACACGGTGGGATCTTCAAGGAGATAGGAAGCCCTGCCGGGAGGATGCTTT 901
Qy      378  ValCysArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrGlyLeu 397
Db      902  GTTTGCAGGCAAAACAGACAGCAGCTGGAGGAGGAGCAGAAAGAAAGCACTGTATGTTTG 961
Qy      398  GluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsn 417
Db      962  GAAGTCGGAGGACGTGGAGGAGTGGCAAGTCTGTGTGGGAAAGTTTCTGCCATCAAT 1021
Qy      418  AlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHis 437
Db      1022  GCCAAACATGTCTCTGTGTTGTCCGGAGCCCCCAGGGGGCTCTCCCCGGCTGCCAC 1081
Qy      438  LeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeu 457
Db      1082  TTGGGAGACGGGTCTTCTGACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTCTG 1141
Qy      458  ArgPheLeuIleArgHisThrAsnGlnAlaAspGlnPheAspPheThrPheValGluVal 477
Db      1142  AGATTTCTCATCAGGCACACCAACAGCAGGACAGTTTGTACTTCTCCTTTTGTGAAGTT 1201
Qy      478  TyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu 497
Db      1202  TATCGCGTCAAGAAATTCAGTTTACGTGCAAGACATGAGGAGATGAGACGAGCGACCTC 1261
Qy      498  LysGluGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCys 517
Db      1262  AAGGAGGGGGGGAAGAGCGCTTTTGGGCACATTTTGCAGCACCCCTCTCTGCTGCTGC 1321

```

```

Qy 518 ThrValSerAanSerSerTrpAanCysAspGlyGluValLeuHisSerProAlaIleGlu 537
Db 1322 ACCGCTCCAAACAGCTCTCTGAACTGCGACGGGAGGTCCTGCACAGCCCTGCCATCGAG 1381

Qy 538 ValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAanProIys 557
Db 1382 GTACAGATCCACTCCCAAGCTGGTTCGACTCTTTCCACGAGGAATGAAGAAATCCGAAG 1441

Qy 558 ProAspSerHisSer 562
Db 1442 CCAGACTCACACAC 1456

RESULT 12
AB079067 2830 bp mRNA linear ROD 27-JUN-2002
LOCUS AB079067 Mus musculus cerk mRNA for ceramide kinases, complete cds.
DEFINITION AB079067
ACCESSION AB079067
VERSION AB079067.1 GI:21624341
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Sugiura, M., Kono, K., Liu, H., Shimizugawa, T., Minekura, H.,
Spiegel, S., and Kohama, T.
Ceramide Kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND
FUNCTIONAL CHARACTERIZATION
J. Biol. Chem. 277 (26), 23294-23300 (2002)
11956206
2 (bases 1 to 2830)
Sugiura, M., Kono, K., Shimizugawa, T., Minekura, H., Spiegel, S. and
Kohama, T.
Direct Submission
Submitted (29-JAN-2002) Masako Sugiura, Sankyo Co., Ltd.,
Pharmacology and Molecular Biology Research Laboratories; 2-58
Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan
(E-mail: msugiura@shina.sankyo.co.jp, Tel:81-3-3492-3131,
Fax:81-3-5436-8565)
FEATURES
Location/Qualifiers
1..2830
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..2830
/gene="cerk"
1..1596
/gene="cerk"
/codon_start=1
/product="ceramide kinases"
/protein_id="BAC01155.1"
/db_xref="GI:21624342"
/translation="MGAMGAEPHLSVLWVKRRCAVLSLEPARALLRWRSPEPGPSA
PGADARVLSVEIIATVEKDDCEKSHSGRWKHNPFATFVRKVRHHRWKWARV
TFWSADQLCHLQTLRLLESLSRPHLLVFINFGKGQGRKIYEKTVAPLPLTL
ASITETIIITHANOAKETLYEINTDSYDGVICVCGDGMFSEVLHGVIGRTQOSAGID
PNHPRVLVPESTLIGIIPAGSTDVCYSTVGTNDAAETSAHHIIIGDPLADVSSVHY
QHTLLRSVLLGCGFVGDLLIKDSEKRWMLVRYDFSGLKTFLSHQYVSGTLSPLPA
QHTVGSPRNDKPCAGGCFVCKQSKQLEEEKALYGLNEAEVEEQVTCGKFLAIN
ATNMSCAPRSPGSLSPFAHLGDSDLILIKCSRNFPLRLIRHTNDDQDFDFTV
EYVRVKFHFSTSKVEDEDNDSDKEQKQKFGKICKDRPSCCTCSASRSSWNGDEVHIS
PAIEVRVHCQLVRLFARGIEES"
ORIGIN
Alignment Scores:
Pred. No.: 8.72e-159 Length: 2830
Score: 2436.50 Matches: 454
Percent Similarity: 90.19% Conservative: 33
Best Local Similarity: 84.07% Mismatches: 48
Query Match: 80.55% Indels: 5
Db: 10 Gaps: 1
US-10-631-958-11 (1-562) x AB079067 (1-2830)

```

```

Qy 26 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValIysGlnArg 45
Db 1 ATGGGGCAATGGGGCGGCGGAGCGCTCACCTCCGTCGTGTGGGTGAAACGCGCAGCGC 60

Qy 46 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 65
Db 61 TGTGCGCTCAGCTGGAGCGCGCGGCTCTGTACGCTGGTGGCGGAGCGCGGAGCGCC 120

Qy 66 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 85
Db 121 GGGCCCTCCGCGCGGTCGATGCCCTCGGTACTGTGTGCGGAGATCATCGCTGTT 180

Qy 86 GluGluThrAspValHisGlyHisGlnGlySerGlyLysTrpGlnLysMetGluLys 105
Db 181 GAGGAAAAGACGACTCGGAGAAACAGCCCTCCAGTGGCGGATGATATAAATGAAAAAT 240

Qy 106 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 125
Db 241 CCGTTCGATTCAAGTCCACCGCTGTGAAGCGAGTACGACATCACCGCTCGAAGTGGCA 300

Qy 126 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 145
Db 301 CGGCTGACCTTCTGGAGCGCGGACGAGCTGTGTACCTGTGGCTGCGACACCTCCGT 360

Qy 146 GluMetLeuGluLysLeuThrSerArgProLysHisLeuValPheIleAasnProPhe 165
Db 361 GGGCTGCTGGAGAGCCCTGACTTCAAGACCGAGACACTTGTGGTATTATCATCAACCTTTC 420

Qy 166 GlyCylLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 185
Db 421 GGAGGGAAGGTGAGGCAAGCGCATCTATCAAAAAACAGTGGCGCTCTGTTTACCTTG 480

Qy 186 AlaSerIleThrThrAspIleIleValThrGluHisAlaIleGlnAlaLysGluThrLeu 205
Db 481 GCTTCCATCACTACGAGATCATCATTTACAGAGCATGCCAACCAAGCAAGAGACATTTA 540

Qy 206 TyrGluIleAasnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 225
Db 541 TACGAGATCAACACAGACAGCTATGATGGCATCGTGTGCGTGTGGGGAGCGCATGTTTC 600

Qy 226 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAan 245
Db 601 AGCGAGTGTGCTATGGGTGCTATTGGGAGGAGCGCAGCAGAGCGCTGTGATACGCCCAAT 660

Qy 246 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 265
Db 661 CACCCCGAGCGGTGCTGGTGCCCGAGTACCTCAGGATCGGCATCATACCCGCGAGGTCC 720

Qy 266 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 285
Db 721 ACAGATTGTGTGTTACTCAACAGTAGGCACAAACACGACAGACATCGGCTTTGCAT 780

Qy 286 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAasnSerThrLeu 305
Db 781 ATCATTTATGGGACTCACTGGCAATAGAGTGTCTCTGTGCACATACCAATAACACGCTG 840

Qy 306 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 325
Db 841 CTGCGGTACTCGGTTTCTCTGCTGGGTACGGTTTCTACGGGGAGACTTAATCAAGACAGC 900

Qy 326 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 345
Db 901 GAAAGAAACGGTGGATGGGCTCGTCGGGTATGATTTCTCAGGGTTGAAGACCTTTCTC 960

Qy 346 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 365
Db 961 TCTCATGACTATGAAGGAGACACTGTCTCTTCTCCAGCAGACAGCACCGTGGGATCT 1020

Qy 366 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 385
Db 1021 CCACGGACAATAAACCTCTGCCGGCTGGGTGCTTCTGTGCGAGCAGACAGCAACACAG 1080

```

QY 386 LeuGluGluGluGluGlyLysAlaLeuTyrGlyLeuGluAlaAalaGluAspValGluGlu 405
Db 1081 CTGGAAGAAGAAAGAAAGCCCTGTATGGCCCTGGGAGAACGCCGAGAGTGGAAAG 1140
QY 406 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 425
Db 1141 TGGCAAGTGACATGTGGGAGTTCTTGCCCATCATGCCACCAACATGTCCTGTGCTGT 1200
QY 426 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 445
Db 1201 CCTCGAGCCCTGGGGGCGTGTCCCAATTTGCCCATCTGGGAGATGGGTCTTCTGACCTC 1260
QY 446 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 465
Db 1261 ATCTTATCCGGAAGTGTCCAGGTTCACATTCCTGAGATTCTCATCCGGACACGAAC 1320
QY 466 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 485
Db 1321 CAGGAGGACCATTCGATTCCTTCTGTTGAGTTTATCGAGTCAAGAAATTCACCTTC 1380
QY 486 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 505
Db 1381 ACGTCCGAAGCAGTGGAGATGAGGACAAATGACTCGAAGGAACAAGAGAGAGAAAGTTT 1440
QY 506 GlyHisLysCysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsn 525
Db 1441 GGAAGATCTTGAAGACAGACAGCCCTTGTGCTTGTGCTCAGGCTCCGGAAGCTCTGGAAC 1500
QY 526 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 545
Db 1501 TCGATGCGGAGATCATGCACAGCCCGCCCATTTGAGGTTCAGGTCCATTCGACGTGCTG 1560
QY 546 ArgLeuPheAlaArgGlyIleGluGlu-----AsnProLysProAspSer 560
Db 1561 CGCTCTTTTGTTCGGGGAATCGAGGAGAGTATAAGCAAGAACCCCAAGGCCAGGAGC 1620

RESULT 13
AK129416
LOCUS Mus musculus mRNA for mKIAA1646 protein. linear ROD 21-NOV-2003
DEFINITION AK129416
ACCESSION AK129416
VERSION AK129416.1 GI:37360495
KEYWORDS FLI_CDNA.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Koseki, H., Hiraoka, S.,
Saga, Y., Nagase, T., Ohara, O. and Koga, H.
Prediction of the coding sequences of mouse homologues of KIAA
gene: III. the complete nucleotide sequences of 500 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
DNA Res. 10 (4), 167-180 (2003)
22977043
MEDLINE
PUBMED 14621295
REFERENCE 2 (bases 1 to 3661)
Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.
Direct Submission
TITLE Submitted (23-JUL-2003) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics, 2-6-7
Kazusa-kanatani, Kisarazu, Chiba 292-0818, Japan
(E-mail: mouse@kazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)
The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing: Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
LOCATION/Qualifiers
1. .3661
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"

gene
CDS
/clone="mph01246"
/tissue_type="embryonic tail"
/note="vector:modified pBC SK+"
1. .3661
/gene="mKIAA1646"
<1..1231
/gene="mKIAA1646"
/note="CDS is predicted by in silico analysis. Start codon
is not identified.
Start codon is not identified."
/codon_start=2
/evidence=not experimental
/product="mKIAA1646 protein"
/protein_id="BAC98226.1"
/db_xref="GI:37360496"
/translation="FCFAASRPKHLVFINPFGKGQKRIYEKTVAPLFTLASITTE
IIITEHANOAKETLYEINTDSYDGI VCVGGDMFSEVLHGVIKTRQDSAGIDPNHPRA
VLVPTLRIGIIPAGSDTCVYSTVGTNDATSAHLIIIGDSLAI DSSVHVHNTLLR
YSVLLGYPFGDLIKDSEKRWMLVRFDSGLKTFLSHQYEGTSLFLPAQHTVGS
PRDNCRAGCFVCRSQQLKEEEKKALYGLNAEVEEVEQVTCGKFLAINTMSC
ACPRSPGLSPFAHLGDGSDLIILIRKSRFNFLRIHTNQDQDFTFVYVRVK
KFHPTSKHVEDENDSKQEKQKFKICKORPSC TCSASRSSMNCDEGVMHSPAIEVR
VHCQLVRLFARGIEES"

ORIGIN
Alignment Scores:
Pred. No.: 1,92e-122 Length: 3661
Score: 1912.50 Matches: 356
Percent Similarity: 92.49% Conservative: 26
Best Local Similarity: 86.20% Mismatches: 26
Query Match: 63.22% Indels: 5
DB: 10 Gaps: 1
US-10-631-958-11 (1-562) x AK129416 (1-3661)
QY 153 SerArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlyGlyGly 172
Db 17 TCAAGACCGAAGCACTTGTGCTGTTATTCATCAACCTTTCCGAGGGAAGGTCAAGGCAAG 76
QY 173 ArgIleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIle 192
Db 77 CGCATCTATGAAAGAACAGTGGCGCTCTGTTTACCTGGCTTCATCAGGAGATC 136
QY 193 IleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLys 212
Db 137 ATCATTCAGAGCATGCCAACCAAGCCAGAGAGACTTTATACGAGATCAACACACAGACAGC 196
QY 213 TyrAspGlyIleValCysValGlyValGlyAspGlyMetPheSerGluValIleHisGlyLeu 232
Db 197 TATGATGGCATCGTGTGCTAGGTGGGACGGCATGTTCCAGGAGGTCTCGATGGGGTG 256
QY 233 IleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuVal 252
Db 257 ATTGGAGGAGCGACAGAGCGCTGTTATCGACCCCAATCACCCCGAGCGGTGCTGTG 316
QY 253 ProSerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTyrSer 272
Db 317 CCCAGTACCTTCAGGATCGGCATCATATCCCGAGGGTCCACAGATTGTGTGTACTTCA 376
QY 273 ThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeu 292
Db 377 ACATGGGCAACAAACGACGACAGACATCGGCTTTGCACATCATTTATGGGAGACTCGCTG 436
QY 293 AlaMetAspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeu 312
Db 437 GCATAGACGTGTCTCTGTGCATACCATTAACACGCTGCTCGGTACTCGGTCTCTCTG 496
QY 313 LeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTyrLeuGly 332
Db 497 CTGGGCTACGGTTTCTACGGGGACTTAATCAAGACAGCGAAGAAACCGGTGATGGGC 556
QY 333 LeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGlyGly 352


```
Db 717 CGATGAGATGAGCTATTTCAGGCTTCAAGACTTTTCTTCTCATCATTCATTATGAAGA 776
Qy 353 ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLeuProCys 372
Db 777 ACAATTTCTTTCAACCAACGAAACACACACTGGGATCTCCACGAGATAAAGATAGCTGC 836
Qy 373 ArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGluGluGlnLysLys 392
Db 837 AGAAGAGATGTTATTTTGAAGGAAAGTGAAGCAACAACTGGCAGAACAAACGAAGAAG 896
Qy 393 AlaLeuTyGlyLeuGluAlaAlaGluAspValGluGluTyrGlnValValCysGlyLys 412
Db 897 TGT-----GGATTCAAACATGAAGAGATGAAGAGATGAAGGTTATTAAAGGGAAA 950
Qy 413 PheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgSerProArgGlyLeu 432
Db 951 TTTCTAGCCATCAATGTAGTAATATGCTGCTGCTCCACGAGTCCAAAGGCTCT 1010
Qy 433 SerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSer 452
Db 1011 TCACGGCAGCTCAATTTAGCAGATGTTTCAGCTGACCTGACCTGATTTAGTTTCGGAATGCTCC 1070
Qy 453 ArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPhe 472
Db 1071 AGATTGATTTCTGCGTATCTTGTCCAGACATACAAACAAAGATGACCAGTTTGACTTC 1130
Qy 473 ThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAsp 492
Db 1131 TCATTTGATGTTTATCGGTTGAAGAGTTTCAATTTACATCAAGCTTTCAGAGAC 1190
Qy 493 GluAspSerAspLeuLysGluGlyCysLysLysArgPheGlyHisIleCysSerSerHis 512
Db 1191 AATGAAGCAGCTGTCACAGATAGAGAAAGAAACATTTTGGCCAGTCTTCAGAGATCAT 1250
Qy 513 ProSerCysCysGlyThrValSerAsnSerSerTyrAsnCysAspGlyGluValLeuHis 532
Db 1251 CCAGCTCTGCTGCAATATTTGCTAATAGCACTTGAATGTTGATGGAGAAACCTTGAT 1310
Qy 533 SerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIle 552
Db 1311 AGCTCAGCAATTGAAGTGAAGGTTTCATGCCAGCTTAATGAACCTGTTTGCAGGGGAATC 1370
Qy 553 GluGluAsnProLysProAsp-----SerHisSer 562
Db 1371 GAGGAACCTGTAAGATGAAGTGCCTGCTACAGCCACAGT 1409

RESULT 15
BC074350 2494 bp mRNA linear VRT 03-AUG-2004
LOCUS Xenopus laevis MGC84197 protein, mRNA (cDNA clone MGC:84197
DEFINITION IMAGE:6954181), complete cds.
ACCESSION BC074350
VERSION BC074350.1 GI:49257664
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 2494)
AUTHORS Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
REFERENCE 2 (bases 1 to 2494)
AUTHORS Klausner,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
```

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, K.J., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.D., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fane, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

3 (bases 1 to 2494)

Klein, S. and Gerhard, D.S.

Direct Submission

Submitted (23-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA

NIH-MGC Project

Contact: XGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement:

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Anara Maseon, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smalilus, Jeff Stott, Miranda Tsai, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 161 Row: a Column: 22

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source

1. 2494

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="MGC:84197 IMAGE:6954181"

/tissue_type="Brain, adult xenopus"

/clone_lib="NICHD XGC_Brn1"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6.1"

1. 2494

/gene="MGC84197"

/db_xref="LocusID:444656"

86. 1804

/gene="MGC84197"

/codon_start=1

/product="MGC84197 protein"

/protein_id="AAH74350.1"

/db_xref="GI:49257665"

/db_xref="LocusID:444656"

/translation="MESAHSTSLAVLCSVLVKRQSCVTLDPGRSLLSWRELPRHG
RESRPGICLTVLHOTGRLVNRTPQPSGGVTPVTEIVSGEATIDEKYNSMKW
OHLCPPAFVTVYQVRKHWRCKEYTFWCSDEQLFYOMLOAFHLLDEQOYRPNL
LVYINPYGKKRGKQIYENKVAFLFSAAIGCADIVITEYAHARDHLYLDANLEKYDGV
VCVGGDGMFSEFVLHGLVIRMQKDSVDHNNPNSAQLSRNCNMRIIPAGSTDCICATYAT

gene

CDS

GINDPETSALHIIILGDCQPLDVCSVHYKRTFLKYSVLLGFGFYGDVLKGSKNRWLG
PARVDVSGFTFLTHHCYEGSVSPQAKWLVGSPRODTCTSGCYICRQSSKOLDEOE
QTOACSEHREDDMTTIKGRFMINAVSMSCACPRSPNGLSPAHLADGSADLIIV
RKSRDLFBLHLIHTSKQDFPFVEYVRVKNFQTPKHFEDENESDIDGLGKK
NFSQICDHFSCGCHNVNSIWNCDGETLDQTAEMRVHQCILKLFARGIENAKLEDIV
NPSKVF"

ORIGIN

Alignment Scores:
Pred. No.: 6.16e-109 Length: 2494
Score: 1714.00 Matches: 322
Percent Similarity: 72.32% Conservative: 83
Best Local Similarity: 57.50% Mismatches: 125
Query Match: 56.66% Indels: 30
DB: 5 Gaps: 4

US-10-631-958-11 (1-562) x BC074350 (1-2494)

```
Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys 42
Db      ::::
Qy 83 GCCATGGAGTCGCTCATTCACCTCGTTAGCCGCTGCTGCTCGGTCTCGGTTAAG 142
Db      ::::
Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer 62
Db      ::::
Qy 143 AGGCAAAAGCTGCGAGGTGACCTCGACCCCGGGGCTCTCTGCTCTCC---TGGAGAGAA 199
Db      ::::
Qy 63 ProGlyProGlyAlaGly-----AlaProGly----- 71
Db      ::::
Qy 200 CTGAGGCCCGACACGGAAGAGAGAGAGCGCGCGCGGGAATCTCTTGACAGAGTGCTC 259
Db      ::::
Qy 72 -----AlaAspAlaCysSer 76
Db      ::::
Qy 260 CACCAAAACAGAGCGCTAGTAAACCGTTTCACTCGGCAATTCCTTCAGGTGGGTACC 319
Db      ::::
Qy 77 ValProValSerGluLeuIleAlaValGluGluThrAspValHisGlyLysHisGlnGly 96
Db      ::::
Qy 320 GTCCAGTGAAGTAAATGTTTCATAGTGGGAGAGAGAGATGATGAGAGTACTACAAAC 379
Db      ::::
Qy 97 SerGlyLysTrpGlnLysMetGluLysProTrpAlaPheThrValHisCysValLysArg 116
Db      ::::
Qy 380 AGTATGAATGGCAGCATATTTGTAAGCCTCGTGCAATTTACAGTACGTTATGTACAGAGA 439
Db      ::::
Qy 117 AlaArgArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeu 136
Db      ::::
Qy 440 GTGAGGAAGCACCGCTGCGGTGCAAGAGAGTGACCTTTTGTGTCAGTGATGACAGCTT 499
Db      ::::
Qy 137 CysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLys 156
Db      ::::
Qy 500 TTTTATCAGTGGTTACAGCAATTTTCATGATCTACTTGAGCAGCAAACTCATAGACCAAAG 559
Db      ::::
Qy 157 HisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTrpGlu 176
Db      ::::
Qy 560 AACTTGCTAGTTTATATTAATCCTTATGTTGGGAGAGAAAGAGGTAAACAGATATATGAG 619
Db      ::::
Qy 177 ArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGlu 196
Db      ::::
Qy 620 AATAAAGTGGCTCCATTATTCAGTGCTGCTGGGATTTGTGCTGATGTATAGTTACTGAG 679
Db      ::::
Qy 197 HisAlaAsnGlnAlaLysGluThrLeuTrpGluLeuAsnIleAspLysTrpAspGlyIle 216
Db      ::::
Qy 680 TATGCAAAACCATGTGAGGATCACCTATATGCAACCTGGAAAAATATGACGGGGTT 739
Db      ::::
Qy 217 ValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThr 236
Db      ::::
Qy 740 GTTTGTGTTGGGGATGGAAATGTTTCAGCGAAGTGTGCAATGGCCTCATTTGTCAGATG 799
Db      ::::
Qy 237 GlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLys 256
Db      ::::
Qy 800 CAAAAAGAGTCTGATGTAGATCACAAACATCCAGTGCACAGTTTATCTCGGTGTAAATAG 859
Db      ::::
Qy 257 ArgIleGlyIleIleProAlaGlySerThrAspCysValCysTrpSerThrValGlyThr 276
Db      ::::
Qy 860 AGGATTGGCATTTATCCCTGCTGTTCAACAGACTGTATATGCTATGCTACTTGTGGCATT 919
Db      ::::
```

```
Qy 277 SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal 296
Db      ::::
Qy 920 AATGACCACAGAGCTGCTGTTTGCACATCATTTCTCGTGACTGCCACGCTCTTGATGTA 979
Db      ::::
Qy 297 SerSerValHisAenSerThrLeuLeuArgTrpSerValSerLeuLeuGlyTrpGly 316
Db      ::::
Qy 980 TGCTCTGTACACTACAAGAGGACATTTCTTAAGTATTTCGGTTTCGTTATTGGGATATGGT 1039
Db      ::::
Qy 317 PheTrpGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTrp 336
Db      ::::
Qy 1040 TTTTATGAGATGTGTTAAAGGCGAGTGAATAAATCGTTGGTTAGTCTCTGCTAGATAT 1099
Db      ::::
Qy 337 AspPheSerGlyLeuLysThrPheLeuSerHisHisCysTrpGluGlyThrValSerPhe 356
Db      ::::
Qy 1100 GACGTTTCAGGTTCAAGACATTTCTGACTCACCATTGCTATGAAGGGTCAGTCTCATTC 1159
Db      ::::
Qy 357 LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 376
Db      ::::
Qy 1160 CAGCCAGCAAAAGTGGGTTCTGGGATCACCAGAGACCAACAACTTGCACATCTGGGTGC 1219
Db      ::::
Qy 377 PheValCysArgGlnSerLysGlnLeuGluGluGluGlnLysLysAlaLeuTrpGly 396
Db      ::::
Qy 1220 TACATTTGCGAGGCAAAAGTAGCAGCAGCTAGACGAGCAAGCAAAACCCAGGCTTGTGGA 1279
Db      ::::
Qy 397 LeuGluAlaAlaGluAspValGluTrpGlnValValCysGlyLysPheLeuAlaIle 416
Db      ::::
Qy 1280 TCAGAACACAGAGAACAGACGATGCTGGCAACAATCAAGGAAGATTTCATGGCAATC 1339
Db      ::::
Qy 417 AsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAla 436
Db      ::::
Qy 1340 AATGCTGTGACGATGAGTGTGTCATGTCCAGAGCCCAATGGCCTTTCACCAGCAGCT 1399
Db      ::::
Qy 437 HisLeuGlyAspGlySerSerAspLeuIleLeuArgLysCysSerArgPheAsnPhe 456
Db      ::::
Qy 1400 CACCTGGCAGATGGAAGTGCAGATCTCATTTTAGTAAAGTGCTCAAGACTTGTGATTTT 1459
Db      ::::
Qy 457 LeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGlu 476
Db      ::::
Qy 1460 TTAAGGCACCTTATCAGGCACACACAGCAATAAAGACCAGTTTGACTTTTCCATTTTGTGAA 1519
Db      ::::
Qy 477 ValTrpArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAsp----- 494
Db      ::::
Qy 1520 GTTTATCGAGTAAAGAAATTTTCAGTTTACTCCAAAACACTTTGAAGATGAAGACAATGAA 1579
Db      ::::
Qy 495 ---SerAspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHisPro 513
Db      ::::
Qy 1580 AGCAGATATAGGAATTTCTTGGCAAGAAAACCTTTTCTCAGATATGCACAGACATCCA 1639
Db      ::::
Qy 514 SerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSer 533
Db      ::::
Qy 1640 TCATGTGGATGAACCATGTGAATAGCATTTTGGAAATTTGTGATGGGAGACATTGGACCAG 1699
Db      ::::
Qy 534 ProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGlu 553
Db      ::::
Qy 1700 ACCGCCATTGAAATGAGGGTCCACTGCGCAACTATATAAATTTATTTGCACGCGGATGAG 1759
Db      ::::
```

Search completed: September 6, 2005, 16:40:11
Job time : 7486.62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 03:56:01 ; Search time 30.7621 Seconds
(without alignments)
1363.781 Million cell updates/sec

Title: US-10-631-958-11
Perfect score: 3025
Sequence: 1 HEANGAPLGVRAAPPART.....QLVRLFARGIENPKPDSSH 562

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap:*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pap:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	801.5	26.5	687	4	US-09-270-767-45874
2	585.5	19.4	359	4	US-09-270-767-46720
3	306	10.1	490	4	US-10-053-510-19
4	302.5	10.0	299	4	US-09-270-767-61405
5	301	10.0	524	4	US-10-053-510-20
6	271.5	9.0	618	4	US-09-970-516-4
7	271.5	9.0	618	4	US-09-817-676A-14
8	267.5	8.8	617	4	US-09-817-676A-12
9	266.5	8.8	384	4	US-09-959-897-2
10	264.5	8.7	384	4	US-09-970-516-2
11	263.5	8.7	384	4	US-09-970-516-1
12	263.5	8.7	384	4	US-09-796-487-3
13	260	8.6	388	4	US-09-817-676A-15
14	260	8.6	388	4	US-09-796-487-2
15	255.5	8.4	368	4	US-10-053-510-21
16	254	8.4	373	4	US-09-796-487-5
17	254	8.4	381	4	US-09-796-487-1
18	254	8.4	381	4	US-09-796-487-4
19	247.5	8.2	382	4	US-09-970-516-6
20	244.5	8.1	392	4	US-09-796-487-6
21	242	8.0	424	4	US-09-796-487-8
22	228	7.5	536	4	US-09-248-796A-15859
23	206.5	6.8	204	4	US-09-796-487-9
24	198	6.5	403	4	US-09-796-487-7
25	161.5	5.3	312	4	US-09-949-016-9811
26	153	5.1	313	4	US-09-107-532A-7154
27	148	4.9	119	4	US-09-205-258-788

28	146	4.8	372	4	US-09-603-208A-264	Sequence 264, App
29	144	4.8	319	4	US-09-270-767-38102	Sequence 38102, A
30	144	4.8	319	4	US-09-270-767-53319	Sequence 53319, A
31	126.5	4.2	293	4	US-09-205-258-328	Sequence 328, App
32	123.5	4.1	304	4	US-09-134-000C-5366	Sequence 5366, App
33	119.5	4.0	283	4	US-09-634-238-373	Sequence 373, App
34	118	3.9	303	4	US-09-107-532A-5133	Sequence 5133, App
35	117	3.9	160	4	US-09-270-767-62339	Sequence 62339, A
36	116.5	3.9	294	4	US-09-583-110-3740	Sequence 3740, App
37	116.5	3.9	315	4	US-09-107-433-3121	Sequence 3121, App
38	113.5	3.8	345	4	US-09-134-000C-3813	Sequence 3813, App
39	110	3.6	324	3	US-09-134-001C-4774	Sequence 4774, App
40	107.5	3.6	325	4	US-09-107-532A-6786	Sequence 6786, App
41	107	3.5	63	4	US-09-959-897-45	Sequence 45, Appl
42	104.5	3.5	64	4	US-09-959-897-37	Sequence 37, Appl
43	103.5	3.4	316	4	US-09-710-279-1350	Sequence 1350, App
44	101.5	3.4	400	4	US-09-252-991A-26145	Sequence 26145, A
45	100	3.3	974	4	US-09-883-134-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-270-767-45874
; Sequence 45874, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45874
; LENGTH: 687
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-45874

Query Match	26.5%	Score	801.5	DB	4	Length	687
Best Local Similarity	31.3%	Pred. No.	9.7e-74				
Matches	197	Conservative	97	Mismatches	206	Indels	129
Gaps	16						
Qy	36	QSVLWVQQRCAVSL	--PARALLRWRS	PGGAGACADACSV	PSVSEIIIAVEETDVHGK	93	
Db	59	QQLWVERLQIKQSPQNEAKPL	-----PPDS	PAQGGICSYGQSHVLHD	--DVVSI	111	
Qy	94	HQSGKWKQMEKP	-----YAF	TVHCVRARR	-----HRWKWAQ	126	
Db	112	RSQDTKASSLKPPSPGSRSSGCGDVAQ	CKPTSOYLTI	NYAMRLSKSQDTCNRWELRR	171		
Qy	127	VTFWCPSEQLCHLWLTREMLEKT	---SRPKLLVFIN	PPFGKGOKRIYERKVALPLF	183		
Db	172	LTFFNSDPYIVRQWDBQLRLHSSSPT	RMRLVFINPY	PGKKGAGAQTERRHRIF	231		
Qy	184	TLASITTDIIIVTEHANAKETLYEIN	DKVDGIVCVG	GDGMFSVHLGLIGRTORSAGVD	243		
Db	232	QLAGVDATCITTQANQVKDILLSHD	LGVDVAVCCV	GDGTVAEVINGLIIFQNRRLGLD	291		
Qy	244	QNHPRVLVPSLRIGIIPAGSTDCVCT	YSTVGTSDAETS	ALHIVVGS	DLAMDVSSVHNS	303	
Db	292	EORPPYIPRP-ALPVGVI	PAGSTDTIAYS	MHTADVTAATHVIL	QHGRLDVCVSNQ	350	
Qy	304	TLLRYSVLLGYGFGDIIKDSEKRW	LGILARYDFS	GLKTFLSHCYEGTVSFL	-----357		
Db	351	SLLFACSVLSYGLGDVAAQSEN	YRWMPRRYSGV	KAFPLNNRGYDAELRMLEEDLL	410		
Qy	358	---PAQITVSGSPRD	-----RKPCR	AGCFVC	-----RQSKQOLEEQKALYGL	398	
Db	411	LTTPLEDIPQSDPSVCSLGRSV	PSVCYANCOR	CSFASSIQEORSLFIQESKEA	----E	466	

```
Qy 399 AAEDVE-----EQVVGCKFLAINAT 419
Db 467 RNQVETEDSHLAASEAALLRPRPGWNLRLPTGSISSMRNLGNDQKVVVRGNFFMICGA 526
Qy 420 NMSCACRRSPGLSPAHLGCGSSDLILIRKCRFNFLRLIRHTNQO-QDFDFTFVEVY 478
Db 527 NITCACARSPNGISRYSHLGGCLDLILVKKTSLLNNVRFLLNTPAGRSGDIRNLPFVEVY 586
Qy 479 RVKKFOFTSKHMEEDSDLKEGGKKRFCHIC-----SSHPSCCCTVSSNWNCDGE 529
Db 587 RTRFRFRFASSEDVSLAGS-----CQPTTPEEMTAHSS-----STEFSSWNCDE 635
Qy 530 VLHSPAIEVRVHCQVRLFARGIENPKP 558
Db 636 VVTDLDTMRSHCQLEIEVFRGPHYSYKP 664

RESULT 2
US-09-270-767-46720
; Sequence 46720, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46720
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46720

Query Match 19.4%; Score 585.5; DB 4; Length 359;
Best Local Similarity 35.5%; Pred. No. 8.7e-52;
Matches 125; Conservative 55; Mismatches 103; Indels 69; Gaps 6;

Qy 154 RPKHLLVFINPFGKGGKRIYERKVAFLFTLASITTDIIVTEHANOAKETLYEINIDKY 213
Db 13 RVRELLVFINPFGKKGAGQYRHRVPIFOLAGVDATCITTRQANQVKDILLSHDIGVY 72
Qy 214 DGIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRVAVLPVSSSLRIGIIPAGSTDVCYST 273
Db 73 DAVCCVGGDGTVAEVINGLIFQRMRELGLDQRPPYIPRP-ALPVGVIAGSTDTIAYSM 131
Qy 274 VGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGYPGYDIIKDSEKKRWGL 333
Db 132 HGTADVRTAAITHVLGQHRGLDVCVSGSGLLFCASVLISGYLGDVAAQSENYRWGP 191
Qy 334 ARYDFSLGKTLFSLHCHYEYGVTFSL-----PAQHTVGSPRD-----RKPCRAGC 376
Db 192 RRYEYGVKAPLNNRGYDAELRMLLEEDPLLTLTTPLEDIPQSDVCSLGSVPSVCYANC 251
Qy 377 FVC-----ROSQOOLEEKKALYCLEAAEDVE-----ERNQVETEDSHLAASEAALLRPRPGNLR 404
Db 252 QRCFSASSIQORSLSLFIQESKEA-----ERNQVETEDSHLAASEAALLRPRPGNLR 307
Qy 405 -----EQVVGCKFLAINATNMSCACRRSPRGLSPAHLGCG 441
Db 308 LPTGSISSMRNLGNDQKVVVRGNFFMICGANITCACARSPNGISRYSHLGDG 359

RESULT 3
US-10-053-510-19
; Sequence 19, Application US/10053510
; Patent No. 6830881
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; APPLICANT: Fyset, Henrik
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; ; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
```

```
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053,510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-053-510-19

Query Match 10.1%; Score 306; DB 4; Length 490;
Best Local Similarity 24.6%; Pred. No. 1.4e-22;
Matches 119; Conservative 68; Mismatches 165; Indels 132; Gaps 19;

Qy 135 QLCHLWLOTLEMLE-KLISRP-----KHLVFINPFGKGGKRIYERKVAFLFTLASIT 189
Db 12 QTAEMMHHTIRKHKRNGSSSPADCGKQLLILNPKSGSGKRELFOKQVAPLLTEAEVQ 71
Qy 190 TDIIVTEHANOAKE--TLYEINIDKYDGVCVCGDMFSEVLHGLIGLIGRTORSAGVDQNH 246
Db 72 YDIQITTHPOYAKEFVTRRDL-LTRYSLIVASGDLGYEVLNGLMER-----MDW 122
Qy 247 PRVAVLPSSLRIGIIPAGSTDVCV-----CYSTVGTSDAETSALHIVVGDLSAMDVSSV 299
Db 123 RRAC---RELPLGIIPCGSGNGLAKSVAAHHCNEPYEPKPIHLHATLTCMACKSTPMDVVRV 179
Qy 300 H---HNSTLLRYSVSLGYPGYDIIKDSEKKRWGLARYDFGLKTLFSLHCHYEYGVTF 356
Db 180 ELATRDKHFMVYSLVSGWGLIADIDIESERLSISGAQRFLLWAIRKLIGLRSYKGRVSY 239
Qy 357 -----LPAQHTVGSPRDRKPCRCAGFCVCRSQKQLEEEKKALYCLEAA 400
Db 240 LLGKGKKEPPVEAARELPRAESTAGIRSSLPLNAGEF-----HDLPEEEGEAVLDGEQFA 295
Qy 401 EDV-----EEQVVC-----GKFLAINATNMSCACR-----RSP 429
Db 296 DAISLDRSVVRQHADSWHSAMSRRTAYISLGGPSMRSNRMSISQRIEANAFAERVP 355
Qy 430 RG-----LSPAHLGDCSSDLILIRK- 450
Db 356 TGTIPPLQMPLLSSDGCWICEDGFVMVHAAATHTLSSDVFFAPESRLDDGLIYLVIIRRG 415
Qy 451 CSRPNFLRLIRHTNQODQF---DFTFVEYVRVKKFOFTSKHMEDESD---LKEGGKKR 504
Db 416 VSRHQLLNFL-NINAGTHLPICGEDPFIVKVPCEAFR-----LEPSSSDGILVVDGERVE 469
Qy 505 FGH 508
Db 470 YGPI 473

RESULT 4
US-09-270-767-61405
; Sequence 61405, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61405
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-61405

Query Match 10.0%; Score 302.5; DB 4; Length 299;
Best Local Similarity 29.2%; Pred. No. 1.5e-22;
```



```
Matches 85; Conservative 35; Mismatches 82; Indels 89; Gaps 9;
QY 342 KTFLEHHCYEGTVSFL-----PAQHTVGSPPD-----RKPCRAGTCVC----- 379
Db 1 KAFLENNRGYDAELRMLEBPDLLLTPTLEDIPQSPDSCVSLGSEVSVYANQRCSPASS 60
QY 380 ---ROSKQOLEEOKKALYGLEAEDVE----- 404
Db 61 IQEQRSSJFIQESKEA-----ERNQVETEDSHLAASEAALLRPRPRGNLRLPTGSISS 116
QY 405 -----EMQVCGKFLAINATMSCARRSPRLGSPAHLGDSGLILIRKCSRNF 457
Db 117 MENLNDQWKVVRGNFFMCGANITACARSPNGISRYSHLGDGCLDLILVKKTSLLNV 176
QY 458 RFLIRHTNQ--DQDPTFVEYVRVKKQFTSKHMEDESDSKLKEGKKRFGHC----- 509
Db 177 RFLNTAGRSGDIRLPLFVEYVTRFRTRPTSAEEDYSLAGS-----CQIPTPPE 228
QY 510 ---SSHPSCCTVSNSSWNCDCGVLHSPAIEVRVHCOLVRLFARGIEENPKP 558
Db 229 ENTAAHSS---STFSSWNCDCGEVVTDLDTMRSHCOLIEVFMRGPHSYSKP 276

RESULT 5
US-10-053-510-20
; Sequence 20, Application US/10053510
; Patent No. 6830881
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; APPLICANT: Fyrest, Henrik
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053.510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-053-510-20

Query Match 10.0%; Score 301; DB 4; Length 524;
Best Local Similarity 25.2%; Pred. No. 5e-22;
Matches 121; Conservative 66; Mismatches 151; Indels 142; Gaps 20;
QY 140 WLQTLREMLEK-----LTSRPKLLVFINPFGKGQKRIYERKVAFLTLASIT 189
Db 17 WYRSLRWQLHRTLEEIFVAPTVDERRRVLNLPKSGSGDAREVFNNMHTVPLNEAEVP 76
QY 190 TDIIVTEHANOAKETLYEINIDKYGIVCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPR 249
Db 77 YDLYTKHNFALIEFLSTFCLDAWCCVAVAGDGLFHEIVNGLLQKQ-----DWAH--- 127
QY 250 VLVPSLSLRIGIIPAGSTDCV-----CYS--TVGTSDAETSALHIVVGDSLAMDVSSVHH 301
Db 128 --VLPHLALGIIPCGSGNGLARSIAHCYNKPVLG-----AALTVISGRSSPMDVVRVQL 179
QY 302 NSTLRYSVSLLYGYFGYGDIIKDSKKRWGLARYDFSLGKTLFSLHHCYEGTVSFLPAQH 361
Db 180 QSRSL-YGFLSIGWGLISDVDIESRIIRMLGQRFVTWTLVRLNLTNYNGRISYLLTDH 238
QY 362 TVGSPRD-----RKPCRACF-----VCRQSKQOLBEE----- 389
Db 239 EVSSTHSATGYAAQRMRQSSRCNTHIDMLNGPAPIYHSSAEYLPQEFADVISLTSINQ 298
QY 390 -----QKALY-----GLEAAB-----DV 403
Db 299 SFRSCDSWLSGSRRSFYYSISEIYHSLADESEFAGLAAASLENRQNYGPASELPDL 358
QY 404 EE-----WQVVCCKFLAINA---TNMSCACRRSPRGLSPAHLGDSGLILIRK-CS 452
```

```
Db 359 NEPLSEDOGWLVEGEFVMMHAYVQTHLIGDCH-----FAPKAQNDGTIYLLIRAGIS 413
QY 453 RFNFLRELIR---HTNQDQDFDTFVEYVRVKKQFTSKHMEDESD---LKEGKKRFG 506
Db 414 RPHLLSFLYNNSSGTHLPESH-DHVKVLPVRAFR-----LEPYDNHGIIITVDGERVEFG 467

RESULT 6
US-09-970-516-4
; Sequence 4, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970.516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-516-4

Query Match 9.0%; Score 271.5; DB 4; Length 618;
Best Local Similarity 30.0%; Pred. No. 7.5e-19;
Matches 78; Conservative 44; Mismatches 111; Indels 27; Gaps 4;
QY 116 RARHRKWAQVTPWCPEEQQLCHLWQLTRMLSEKLTSPKHLVFINPFGKGQKRIY 175
Db 111 RAEARWATAITCL-----LRGLPFGDGSITPDLLPRPRLLLVNPFQGRGLAWQMC 164
QY 176 ERKVAPLTLASITTDIIVTEHANOAKETLYEINIDKYGIVCVGGDGMFSEVLHGLIGR 235
Db 165 KHVPLMISEAGLSFNLIQTERQHARELVQGLSEWDGIVTVSGDGLLHEVLNGLLDR 224
QY 236 TORSAGVDQNHPRAVLPFSSLRIGIIPAGSTDCVCYS-----TVGTSDAETSALHI 286
Db 225 -----PDWEEAVKMP-----VGLPCGSGNALAGAVNQHGPFEPALGLDLLNCSLLL 272
QY 287 VVGSLSAMDVSSVHHNSTLLRYSVSLLYGYFGYGDIIKDSKKRWGLARYDFSGKLTFLS 346
Db 273 CRGGHPLDLLSVTLASGRCFSLVAVGFWSDVDIQSRFALGSARFTLTGTVLGLAT 332
QY 347 HHCYEGTVSFLPAQHTVGSPP 366
Db 333 LHTYGRSLYLPATVPEPASP 352

RESULT 7
US-09-817-676A-14
; Sequence 14, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817.676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 14
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-817-676A-14
```


; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases

; FILE REFERENCE: 4-31617

; CURRENT APPLICATION NUMBER: US/09/970,516

; CURRENT FILING DATE: 2001-10-04

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 384

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-970-516-2

Query Match 8.7%; Score 264.5; DB 4; Length 384;

Best Local Similarity 25.3%; Pred. No. 1.9e-18;

Matches 107; Conservative 68; Mismatches 169; Indels 79; Gaps 16;

Qy 154 RPKHLLFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKY 213

Db 13 RPCRVLVLLNPRGGKGKALQFRSHVQPLLAEEAISFTLMLTERNHARELVRSSEELGRW 72

Qy 214 DGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRTIGIIPAGSTDCVC--- 270

Db 73 DALVVMGSDGLMHVVGMLMERPDWETAIQK-----PLCSLPAGSGNALAASL 120

Qy 271 -----YSTVGTSDAETSALHIVVGDSLA-MDVSSVHHNSTLLRYSVSLGFGFYCDI1KD 324

Db 121 NHYAGEYQVTNEDLLTNTCLLRCRLSPMNLLSLHTASGLRFLSVLSLAWGFIADVDLE 180

Qy 325 SEKKRWGLARYDFSLGTLFSLHSHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQSKQ 384

Db 181 SEKYRRLGEMRFTLTGTLRLAALTYRGRLAYLPVGR-VGSKTPASP-----VVVQGPV 234

Qy 385 QLEBEQKALYGLEAAEDV-BEWOVVCCK-----FLAINATNMSCACRRSPRGLSPAHLG 439

Db 235 DAH-----LVPLE--EPVPSHWTVPDDEDFVLVLLALLSHLGSEMFAAPMGRCAC 282

Qy 440 DGSSDLILIRK-CSRFNPLR-FLIRHTNQOQDFTFVYVYRVKKFQFTSKHMEDESD 497

Db 283 -GVMHLFVYRAGVSRAMLLRFLAMEKGRHM-----EYECPLYVYVVAFR-----LEPK 333

Qy 498 KEGGKRFKHCSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENP- 556

Db 334 ---GKGVFA-----VDGELMWSEAVQGVHPNYFMVWGCVPEPPS 371

Qy 557 -KP 558

Db 372 WKP 374

RESULT 11

US-09-949-016-7026

; Sequence 7026, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7026

; LENGTH: 384

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-7026

Query Match 8.7%; Score 263.5; DB 4; Length 384;

Best Local Similarity 25.1%; Pred. No. 2.4e-18;

Matches 107; Conservative 69; Mismatches 164; Indels 87; Gaps 17;

Qy 154 RPKHLLFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKY 213

Db 13 RPCRVLVLLNPRGGKGKALQFRSHVQPLLAEEAISFTLMLTERNHARELVRSSEELGRW 72

Qy 214 DGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRTIGIIPAGSTDCVC--- 270

Db 73 DALVVMGSDGLMHVVGMLMERPDWETAIQK-----PLCSLPAGSGNALAASL 120

Qy 271 -----YSTVGTSDAETSALHIVVGDSLA-MDVSSVHHNSTLLRYSVSLGFGFYCDI1KD 324

Db 121 NHYAGEYQVTNEDLLTNTCLLRCRLSPMNLLSLHTASGLRFLSVLSLAWGFIADVDLE 180

Qy 325 SEKKRWGLARYDFSLGTLFSLHSHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQSKQ 384

Db 181 SEKYRRLGEMRFTLTGTLRLAALTYRGRLAYLPVGR-VGSKTPASP-----VVVQGPV 234

Qy 385 QLEBEQKALYGLEAAEDV-BEWOVVCCK-----FLAINATNMSCACRRSPRGLSPAHLG 439

Db 235 DAH-----LVPLE--EPVPSHWTVPDDEDFVLVLLALLSHLGSEMFAAPMGRCAC 282

Qy 440 DGSSDLILIRK-CSRFNPLR-FLIRHTNQOQDFTFVYVYRVKKFQFTSKHMEDE 493

Db 283 -GVMHLFVYRAGVSRAMLLRFLAMEKGRHM-----EYECPLYVYVVAFR-----LEPK 332

Qy 494 DSDLKEGKRFKHCSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIE 553

Db 333 D-----GKGVFA-----VDGELMWSEAVQGVHPNYFMVWGCV 367

Qy 554 ENP--KP 558

Db 368 PPSWKP 374

RESULT 12

US-09-796-487-3

; Sequence 3, Application US/09796487

; Patent No. 6830916

; GENERAL INFORMATION:

; APPLICANT: Spiegel, Sarah

; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use

; FILE REFERENCE: 07320001aa (2033957-0001)

; CURRENT APPLICATION NUMBER: US/09/796,487

; CURRENT FILING DATE: 2001-03-02

; PRIOR APPLICATION NUMBER: US 60/186,532

; PRIOR FILING DATE: 2000-03-03

; PRIOR APPLICATION NUMBER: US 09/530,868

; PRIOR FILING DATE: 2000-05-05

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 384

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1)..(384)

; OTHER INFORMATION: SEQ ID NO 3 is the peptide sequence of hSPHK1 in Fig. 3, correspo

; OTHER INFORMATION: nding to amino acid residue 1 to 384 of Homo sapiens SPHK-1 of Ge

; OTHER INFORMATION: nBank sequence Accession Number AAF73423.

; PUBLICATION INFORMATION:

; AUTHORS: Nava et al

; TITLE: Functional characterization of human sphingosine kinase-1

; JOURNAL: FEBS Lett.

; VOLUME: 473

; ISSUE: 1

; PAGES: 81-84

; DATE: 2000

; DATABASE ACCESSION NUMBER: AAF73423

```
; DATABASE ENTRY DATE: 2000-06-01
; RELEVANT RESIDUES: (1)...(384)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF73423
; DATABASE ENTRY DATE: 2000-06-01
; RELEVANT RESIDUES: (1)...(384)
US-09-796-487-3

Query Match      8.7%; Score 263.5; DB 4; Length 384;
Best Local Similarity 25.1%; Pred. No. 2.4e-18;
Matches 107; Conservative 69; Mismatches 164; Indels 87; Gaps 17;

Qy 154 RPKHLVFINPFGGKGKRIYERKVAFLFTLASITDDIIIVTEHANOAKETLYEINIDKY 213
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 13 RPRVLVLLNPRGGKQALQFRSHVQPLLAEEISFTLMLTERRNHARELVRSEELGRW 72

Qy 214 DGIYCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGSTDVCV--- 270
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 73 DALVMSGDGLMHVEVNVGLMERPDWETAIQK-----PLCSLPAGSGNALAASL 120

Qy 271 -----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHNSTLLRYSVSLGFGYVDI1KD 324
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 NHYAGYEQVTNEDLLTNTCLLCRRLLSPMNLISHTASGLRFLSVLSLAWGFADVDLE 180

Qy 325 SEKKRWGLARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFYCROSKQ 384
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 SEKYRRLGEMRFTLTGTLRLAALRTYRGLAYLPVGR-VGSKTPASP-----VVVQGPV 234

Qy 385 QLEEQKALYGLAAEDV-EWQVVCVK-----FLAINTNMSCACRRSPRGLSPAHLG 439
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 235 DAH-----LVPLE--EPVPSHWTMPDDEFLVILALLHSLGSEMPAAMPGRCAA----- 282

Qy 440 DGSSDLILIRK-CSRFNPLRLI-----RHTNQDQDFTFVEYVRVKFQFTSKHMEDE 493
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 283 -GVNHLFYVRAGVSRAMLLRFLANEKGRM-----EYECPLVTVVPVAFR-----LEPK 332

Qy 494 DSDLKEGKGRFGHICSHSPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIE 553
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 333 D-----GKGVFA-----VDGELMVSEAVQGVHPNYFMVMVSGCIVE 367

Qy 554 ENP--KP 558
    |||
Db 368 PPSWKP 374

RESULT 13
US-09-817-676A-15
; Sequence 15, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; PRIOR FILING DATE: 2001-03-26
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of
; TITLE: murine sphingosine kinase
; JOURNAL: J. Biol. Chem.
; VOLUME: 273
; ISSUE: 37
; PAGES: 23722-23728
; DATE: 1998-09-11
```

```
; DATABASE ACCESSION NUMBER: AAC61698
; DATABASE ENTRY DATE: 1998-09-26
US-09-817-676A-15

Query Match      8.6%; Score 260; DB 4; Length 388;
Best Local Similarity 23.3%; Pred. No. 5.6e-18;
Matches 100; Conservative 70; Mismatches 173; Indels 86; Gaps 13;

Qy 129 FWCPEEQCHLWLTQRLMLEKLTSPKHLVFINPFGGKGKRIYERKVAFLFTLASI 188
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 WWC-----CVLFVV---ECPRGLLPRCRVLVLLNPOGGKQALQLFQSRVQPFLEAEI 53

Qy 189 TTDIIIVTEHANOAKETLYEINIDKYIGIVCVGGDGMFSEVLHGLIGRTORSAGVDQNHPR 248
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 54 TFXKLILTERKNHARELVCAEELGHWDALAVMSGDGLMHVEVNVGLMERPDWETAIQK--- 109

Qy 249 AVLVPSSLRIGIIPAGSTDVCV-----YSTVGTSDAETSALHIVVGDLSA-MDVSSV 299
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 -----PLCSLPAGSGNALAASVNHYAGYEQVTNEDLLINCTLLLCRRRLSPMNLISL 161

Qy 300 HNSTLLRYSVSLGFGYVDI1KDSEKKRWGLARYDFSLGKTLFSLHHCYEGTVSFLPA 359
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 162 HTASGLRLYSVLSLWGFADVDLESEKYRRLREIRFTVGTFFRLASRLRYQGLAYLP- 220

Qy 360 QHTVGSPRDRKPCRAGCFYCROSKQLEEQKALYGLAAEDV-EWQVVCVK-----FL 414
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 221 ---VGTVASKRPA-----STLVQKGPVDTHLVPLEEPVPVSHWTVPVEQDFVLVL 266

Qy 415 AINATNMSCACRRSPRGLSPAHLGDSSDLILIRK-CSRFNPLR-FLIRHTNQDQDFD 472
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 267 VLLHHTLSSLELFAAMPGRCEA-----GVNHLFYVRAGVSRALLRLFLAMQKGKHMELDC 321

Qy 473 TFVEYVRVKFQFTSKHMEDESDLEKGGKGRFGHICSHSPSCCCTVSNSSWNCDEVLH 532
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 322 PYLVHPVPAFRLEPRS-----ORGVFSDGELMV 351

Qy 533 SPAIEVRVH 541
    |||:|||||
Db 352 CEAVQGVH 360

RESULT 14
US-09-796-487-2
; Sequence 2, Application US/09796487
; Patent No. 6830916
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(388)
; OTHER INFORMATION: SEQ ID NO 2 is the peptide sequence of SPHK1b in Fig. 1, correspo
; OTHER INFORMATION: nding to amino acid residue 1 to 388 of SPHK1b of GenBank sequenc
; OTHER INFORMATION: e Accession Number AAC61698.
; PUBLICATION INFORMATION:
; AUTHORS: Kohama et al.
; TITLE: Molecular cloning and functional characterization of murine sphingosine
; TITLE: kinase
; JOURNAL: Journal of Biological Chemistry
; VOLUME: 273
```

```
; ISSUE: 37
; PAGES: 23722-23728
; DATE: 1998
; DATABASE ACCESSION NUMBER: AAC61698
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (1) (388)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61698
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (1) (388)
US-09-796-487-2

Query Match      8.6%; Score 260; DB 4; Length 388;
Best Local Similarity 23.3%; Pred. No. 5.6e-18;
Matches 100; Conservative 70; Mismatches 173; Indels 86; Gaps 13;

Qy 129 FWCPEQLCHLWQLTREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASI 188
Db 2 WWC-----CVLFVV---ECPRGLLPKPCVRLVLLNPQGGKALQLFQSRVQPFLEAEI 53

Qy 189 TTDIIIVTEHANOAKETLYEINIDKYDGIYCVGGDMGFSEVLHGLIGRTQRSAGVDQNHPR 248
Db 54 TFKLILTERKNHARELVCAEELGHWDALAVMSGDGLMHEVVNGLMERPDWETAIQK---- 109

Qy 249 AVLVPSSLRIGIIPAGSTDCVC-----YSTVCTSDAETSALHIVVGDLSA-MDVSSV 299
Db 110 -----PLCSLPGSGNALAASVNHVAGYEQVTNEDLLINCTLLCCRRLLSPMNLISL 161

Qy 300 HNNSTLLRYSVLLGYGFVGDIIKDSKKRWGLGLARYDFSLGKTFSLHHCEGTVSFLPA 359
Db 162 HTASGLRILYSLWSGFWADVLDSEKYRLGEIRFTVGTFFRLASIRIYQGQLAYLP- 220

Qy 360 QHTVGSPRDRKPCRAGCFVCRQSKQOLEEBEOKKALYGLEAAEDV-BEWOVVCVK-FL 414
Db 221 ---VGTASKRPA-----STLVQKGVVDTHLVPLEBPVPSHWTVPVEQDFVLVL 266

Qy 415 AINATNMSCACRRSPRGLSPAHLGDGSSDLILIRK-CSRFNPLR-FLIRHTNQDQDF 472
Db 267 VLLHTHLSSELFAAPMGRCIA-----GYMHLFYVRAGVSRALRLFLAMQKGMELDC 321

Qy 473 TFVEYVRVKFKFTSKHMEDESDLKEGKKRFGHICSSHPSCCCTVSNSSWNCDEVLH 532
Db 322 PYLHVFPVAFRLPR-----ORGVFSDGELMV 351

Qy 533 SPAIEVRVH 541
Db 352 CEAVQGVQH 360
```

```
RESULT 15
US-10-053-510-21
; Sequence 21, Application US/10053510
; Patent No. 6830891
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053,510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-053-510-21
```

```
Query Match      8.4%; Score 255.5; DB 4; Length 368;
Best Local Similarity 25.1%; Pred. No. 1.5e-17;
Matches 105; Conservative 68; Mismatches 167; Indels 79; Gaps 16;
```

```
Qy 158 LLVFNPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKYDGI 217
Db 1 VLVLLNPRGGKALQLFRSHVQPLLAETISFTMLTERRNHARELVSRSELGWDALV 60

Qy 218 CVGGDMGFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVC----- 270
Db 61 VMGSGLMHEVVNGLMERPDWETAIQK-----PLCSLPGSGNALAASLHNYA 108

Qy 271 -YSTVCTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLIGYGFYGDIIKDSKK 328
Db 109 GYEQVTNEDLLTNCCTLLCCRRLLSPMNLISLHTASGLRFLFSVLSLANGFIADVLESEKY 168

Qy 329 RWLGLARYDFSGKTFSLHHCEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSKQOLEE 388
Db 169 RRLGEMRFTLTGTLRLAALRTYRGLAYLPVGR-VGSKTTPASP-----VVYQGVDAH- 221

Qy 389 EOKKALYGLEAAEDV-BEWOVVCVK-----FLAINATNMSCACRRSPRGLSPAHLGDGSS 443
Db 222 -----LVPLE--BPVPSHWTVPDEDFVLVLLHSHLGSEMFAAPMGRCIA--GVM 269

Qy 444 DLILIRK-CSRFNPLR-FLIRHTNQDQDFTFVEYVRVKFKFTSKHMEDESDLKEGG 501
Db 270 HLFYVRAGVSRAMLLRLFLAMEKGRHMEYECPLYLVVYPVVAFR-----LEPKD-----G 318

Qy 502 KKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHLCQLVRLFPARGIEENP--KP 558
Db 319 KGVPA-----VDGELMVSEAVQGVHPHNYFWMVSGCVPEPPPSWKP 358
```

Search completed: September 3, 2005, 04:10:38
Job time : 32.7621 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:09:18 ; Search time 114.766 Seconds
(without alignments)
1893.930 Million cell updates/sec

Title: US-10-631-958-11
Perfect score: 3025
Sequence: 1 HEAANGPAPLGVRAAPPWRT.....QLVRLFARGIENKPDQSHS 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3025	100.0	562	ABB07857	Abb07857 Human sph
2	2888	95.5	537	ABB07856	Abb07856 Human sph
3	2888	95.5	537	ADA05680	Ada05680 Human NOV
4	2888	95.5	537	ADJ96664	Adj96664 Human lip
5	2888	95.5	537	ADN62845	Adn62845 Human NOV
6	2880	95.2	537	AAW49115	Aaw49115 Human cer
7	2588	85.6	481	ABR56302	Abx56302 Human sph
8	2463	81.4	460	AAI96059	Aay96059 Human sph
9	2456.5	81.2	471	AAE07884	Aae07884 Human sph
10	2428	80.3	531	ADP55248	Adp55248 Human PRO
11	2350	77.7	536	ADs11054	Adsl1054 Human the
12	2215	73.2	746	ABG13541	Abg13541 Human hum
13	2210	73.1	416	ABR56301	Abx56301 Human sph
14	1714.5	56.7	727	ABG13543	Abg13543 Novel hum
15	1640.5	54.2	326	ABB07854	Abb07854 Human sph
16	1361.5	45.0	454	ABG13544	Abg13544 Novel hum
17	1159	38.3	255	ADs12267	Adsl2267 Human the
18	1055	34.9	228	AA841822	Aab41822 Human ORP
19	1032	34.1	190	ABP64913	Abp64913 Human pro
20	986	32.6	182	AA842383	Aab42383 Human ORF
21	675	22.3	596	ABB69669	Abb69669 Drosophil
22	664	22.0	136	ABG13540	Abg13540 Novel hum
23	647	21.4	144	AAE07885	Aae07885 Partial r
24	628	20.8	136	ABG13542	Abg13542 Novel hum
25	600	19.8	532	ADQ88891	Adq88891 Novel hum

ALIGNMENTS

RESULT 1

ABB07857

ID ABB07857 standard; protein; 562 AA.

AC ABB07857;

DT 03-JUL-2002 (first entry)

DE Human sphingosine kinase-like protein.

KW Human sphingosine kinase-like protein; intracellular signalling;

KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;

KW autoimmune disease; rheumatoid arthritis; Parkinson's disease.

OS Homo sapiens.

PN WO200228906-A2.

XX 11-APR-2002.

PF 05-OCT-2001; 2001WO-EP011516.

PR 06-OCT-2000; 2000US-0238005P.

PR 23-AUG-2001; 2001US-0314113P.

XX (FARB) BAYER AG.

XX Kossida S, Encinas J;

XX WPI; 2002-340094/37.

XX N-PSDB; ABL40828.

New reagent for modulating the activity of sphingosine kinase-like protein polypeptide or polynucleotide and treating cancer, asthma, allergy, an autoimmune disease, or a central or peripheral nervous system disorder.

Claim 25; Fig 11; 120pp; English.

The invention relates to a human sphingosine kinase-like protein. The polypeptide can be expressed by standard recombinant methodology. The sphingosine kinase-like protein and gene can be used to regulate intracellular signalling and consequently cell proliferation and apoptosis. Such regulation is useful for treating cancer, allergies (e.g. asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and peripheral nervous system disorders (e.g. Parkinson's disease). The present sequence represents the human sphingosine kinase-like protein

DE Human NOV9a protein SEQ ID NO:40.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

KW immunomodulator; cytosatic; nootropic; neuroprotective;

KW antiparkinsonian; antilipaeamic; gene therapy; human disease;

KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX

OS Homo sapiens.

XX

XX WO2003029424-A2.

XX

XX 10-APR-2003.

XX

XX 02-OCT-2002; 2002WO-US031373.

XX

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 19-APR-2002; 2002US-0373884P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.

XX

PA (CURA-) CURAGEN CORP.

XX

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;

PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX

XX WPI; 2003-381626/36.

DR N-PSDB; ADA05679.

XX

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,

PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,

PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX

XX Claim 1; Page 135; 586pp; English.

XX

XX The present invention describes NOVX proteins, where X can be 1 to 55

CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide

CC described above and a carrier; (2) a kit comprising, in one or more

CC containers, the composition described above; (3) an isolated nucleic acid

CC molecule which encodes a NOVX protein of the invention; (4) a vector

CC comprising the nucleic acid molecule described above; (5) a cell

CC

CC comprising the above vector; (6) an antibody that immunospecifically

CC binds to the polypeptide described above; (7) methods for determining the

CC presence or amount of the above polypeptide or nucleic acid molecule in a

CC sample; (8) methods for determining the presence of or predisposition to

CC a disease associated with altered levels of expression of the above

CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a

CC method of identifying an agent that binds to the polypeptide described

CC above; (10) a method for identifying a potential therapeutic agent for

CC use in treating a pathology that is related to an aberrant expression or

CC aberrant physiological interactions of the polypeptide; (11) a method of

CC screening for a modulator of activity or of latency or predisposition to

CC a pathology associated with the polypeptide; (12) a method for modulating

CC the activity of the polypeptide described above; (13) methods of treating

CC or preventing a pathology associated with the above polypeptide in a

CC mammal; and (14) a method for producing the above polypeptide. NOVX

CC sequences have antidiabetic, anorectic, antibacterial, virucide,

CC immunomodulator, cytosatic, nootropic, neuroprotective, antiparkinsonian

CC and antilipaeamic activities, and can be used in gene therapy. The

CC polypeptide is useful in manufacturing a medicament for treating a

CC syndrome associated with a human disease. The polypeptide or the nucleic

CC acid molecule may be used to diagnose, treat or prevent metabolic

CC disorders such as diabetes or obesity, infections, cachexia, cancer,

CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's

CC disease, immune disorders, haematopoietic disorders and various

CC dyslipidaemias. The nucleic acids can also be used as hybridisation

CC probes, in chromosome mapping, tissue typing, preventive medicine and

CC pharmacogenomics. The present sequence represents a human NOVX from the

CC present invention.

XX

XX Sequence 537 AA;

XX

Query Match 95.5%; Score 2888; DB 6; Length 537;

Best Local Similarity 100.0%; Pred. No. 1.3e-269;

Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAPLQSLVWVQQRCAVSLPAPARALLRWRSFPGAGAPGADACVSPVSEIIAV 85

Db 1 MGATGAAPLQSLVWVQQRCAVSLPAPARALLRWRSFPGAGAPGADACVSPVSEIIAV 60

Qy 86 EETDVHGKHGSGKWKQMEKPYAFTVHCVKRARRHWKWAQVTFWCPEEQQLCHLWQLTLR 145

Db 61 EETDVHGKHGSGKWKQMEKPYAFTVHCVKRARRHWKWAQVTFWCPEEQQLCHLWQLTLR 120

Qy 146 EMLEKLTSRPKHLLVFNPFGKGQGRKIYERKVAFLFTLASITTDIIIVTHANQAKETL 205

Db 121 EMLEKLTSRPKHLLVFNPFGKGQGRKIYERKVAFLFTLASITTDIIIVTHANQAKETL 180

Qy 206 YEINIDKYDGIVCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAGS 265

Db 181 YEINIDKYDGIVCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAGS 240

Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGAGFYGDIIKDS 325

Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGAGFYGDIIKDS 300

Qy 326 EKKRWLGLARYDFSGSLKTLFSLHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQSQKQ 385

Db 301 EKKRWLGLARYDFSGSLKTLFSLHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQSQKQ 360

Qy 386 LEEBQKALYGLEAAEDVEEMQVCGKFLATNATNMSCACRRSPRGLSPAHLGSSDL 445

Db 361 LEEBQKALYGLEAAEDVEEMQVCGKFLATNATNMSCACRRSPRGLSPAHLGSSDL 420

Qy 446 ILIRKCSRFNPLRFLIRHTNQDQDFTFVEVYRVKXQFTSKHMEDESDSLKEGGKGRF 505

Db 421 ILIRKCSRFNPLRFLIRHTNQDQDFTFVEVYRVKXQFTSKHMEDESDSLKEGGKGRF 480

Qy 506 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLPARGIEENPKPDSHS 562

Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLPARGIEENPKPDSHS 537

RESULT 4

PR 28-MAY-2002; 2002US-0383656P.
PR 29-MAY-2002; 2002US-0383831P.
PR 25-JUN-2002; 2002US-0391335P.
XX
PA (SMIT/) SMITHSON G.
PA (MILL/) MILLET I.
PA (PEYM/) PEYMAN J A.
PA (KEKU/) KEKUDA R.
PA (JUJU/) JU J.
PA (LILL/) LI L.
PA (GUOX/) GUO X.
PA (PATT/) PATTURAJAN M.
PA (SPYT/) SPYTEK K A.
PA (EDIN/) EDINGER S R.
PA (ELLE/) ELLERMAN K.
PA (MALY/) MALYANKAR U M.
PA (ORTT/) ORT T.
PA (GORM/) GORMAN L.
PA (ZRHU/) ZERHUSEN B D.
PA (ANDE/) ANDERSON D W.
PA (ZHON/) ZHONG M.
PA (CATT/) CATTERTON E.
PA (JIWW/) JI W.
PA (MILL/) MILLER C E.
PA (RAST/) RASTELLI L.
PA (STON/) STONE D J.
PA (PENA/) PENA C E A.
PA (SHEN/) SHENOY S G.
PA (SHIM/) SHIMKETS R A.
PA (ROTH/) ROTHENBERG M E.
PA (LEAC/) LEACH M D.
PA (AGEE/) AGEE M L.
PA (BERG/) BERGHS C.
PA (DIPJ/) DIPIPPO V A.
PA (EISE/) EISEN A.
PA (GANG/) GANGOLLI E A.
PA (RIEG/) RIEGER D K.
PA (SPAD/) SPADERNA S K.
XX
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patturajan M, Spytke KA, Edinger SR, Ellerman K, Malyankar UM;
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI Shmkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;
XX
XX MPI; 2004-213931/20.
DR N-PSDB; ADN62844.
XX
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.
XX
PS Claim 1; SEQ ID NO 40; 395pp; English.
XX
CC The invention relates to isolated NOVX polypeptides and polynucleotides.
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or
CC treat a medical condition in human related to the aberrant expression and
CC activity of NOVX polypeptides. For example, NOVX polypeptides and
CC polynucleotides may be used to treat disorders associated with decreased
CC expression or activity of NOVX by supplementing the patient our
CC production or to rectify mutations. Conversely, antisense NA molecules
CC may be administered to down regulate expression of NOVX polypeptides by
CC binding with the cells own genes and preventing their expression. NOVX
CC polynucleotides and complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar sequences in samples, and so which patients may be in need of
CC restorative therapy. NOVX polypeptides may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of NOVX. The
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be
CC used to modulate NOVX polynucleotide expression and activity of NOVX
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX

CC polypeptides and polynucleotides may be used in this way to prevent,
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
CC haematopoietic disorders, and the various dyslipidaemias, metabolic
CC disturbances associated with obesity, the metabolic syndrome X and
CC wasting disorders associated with chronic diseases and various cancers.
CC They may also be used as antibacterial agents. The present sequence
CC represents the amino acid sequence of a human NOVX protein.
XX
SQ Sequence 537 AA;
Query Match 95.5%; Score 2888; DB 8; Length 537;
Best Local Similarity 100.0%; Prod. No. 1.3e-269;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 MGATGAAPLQSLVWVKQRCVAVSLPARALLRWRSPGCGAGAPGADACSVPYSEIIV 85
DB 1 MGATGAAPLQSLVWVKQRCVAVSLPARALLRWRSPGCGAGAPGADACSVPYSEIIV 60
QY 86 EETDVHGKQSGKQWKQKPYAFTVHCVRARRHRWKQAQVTFWCPPEQLCHLWQLTLR 145
DB 61 EETDVHGKQSGKQWKQKPYAFTVHCVRARRHRWKQAQVTFWCPPEQLCHLWQLTLR 120
QY 146 EMLEKLTSPRKHLVFINPFGKGQGRIVRYKVAFLFTLASITTDIIVTSHANQAKETL 205
DB 121 EMLEKLTSPRKHLVFINPFGKGQGRIVRYKVAFLFTLASITTDIIVTSHANQAKETL 180
QY 206 YEINIDKYGIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSURIGIIPAGS 265
DB 181 YEINIDKYGIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSURIGIIPAGS 240
QY 266 TDCVCTVTGTSDAETSAHIVVGDLSAMDVSSVHHNSTLLRYSVLLGYGYGDIKDS 325
DB 241 TDCVCTVTGTSDAETSAHIVVGDLSAMDVSSVHHNSTLLRYSVLLGYGYGDIKDS 300
QY 326 EKKRWLGLARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPPDRKPCRAGFCVCRSQKQ 385
DB 301 EKKRWLGLARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPPDRKPCRAGFCVCRSQKQ 360
QY 386 LEEBOKKALVLEAAEDVEWVQVCGFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445
DB 361 LEEBOKKALVLEAAEDVEWVQVCGFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
QY 446 ILIRKCSRFLRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLEKGGKKRF 505
DB 421 ILIRKCSRFLRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLEKGGKKRF 480
QY 506 GHICSSHPSCCTVSNSSWNCDEVLHSPAEIVRVHCOLVRLFARGIEENPKPDSSH 562
DB 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAEIVRVHCOLVRLFARGIEENPKPDSSH 537
RESULT 6
AAM49115
ID AAM49115 standard; protein; 537 AA.
XX
AC AAM49115;
XX
DT 20-MAY-2002 (first entry)
XX
XX Human ceramide kinase hCERK1.
KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;
KW neurological disease; inflammation; human immunodeficiency virus;
KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;
KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;
KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cyostatic;
KW enzyme.
XX
OS Homo sapiens.
XX
PN W0200196575-A1.


```

Db      361  SDDLILIRKCSRNFRLIRTNQDQDFTFVEVYRVKQFTSKMEDESDLKEGG 420
QY      502  KKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDH 561
Db      421  KKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDH 480
QY      562  S 562
Db      481  S 481

RESULT 8
AAV96059
ID  AAV96059 standard; protein; 460 AA.
AC  AAV96059;
XX
XX  05-DEC-2000 (first entry)
XX  Human sphingosine kinase C.
XX
XX  Sphingosine kinase C; SKC; human; drug screening; infection;
XX  antiinflammatory; antiallergic; anticancer; inflammation; allergy;
XX  cancer; therapy; diagnosis.
XX
XX  Homo sapiens.
XX
XX  WO200052173-A2.
XX
XX  08-SEP-2000.
XX
XX  02-MAR-2000; 2000WO-CA000223.
XX
XX  02-MAR-1999; 99US-0122516P.
XX
XX  (ALLX ) NPS ALLELIX CORP.
XX
XX  Munroe D, Gupta A, Falzone GR;
XX
XX  WPI; 2000-572185/53.
XX  N-PSDB; AAA50510.
XX
XX  New human sphingosine kinase A, B and C polynucleotides and polypeptides
XX  useful in e.g. chromosome and gene mapping, and detecting inflammation or
XX  disease associated with abnormal levels of sphingosine kinase expression.
XX
XX  Disclosure; Fig 9; 81pp; English.
XX
XX  The present sequence is that of human sphingosine kinase C (SKC), an
XX  enzyme that phosphorylates sphingosine to form sphingosine 1-phosphate.
XX  The sequence was deduced from that of a polynucleotide (see AAA50510)
XX  isolated from an HeLa cDNA library. The invention provides
XX  polynucleotides (see AAA50508-10) and polypeptides (see AAY96057-59) for
XX  the human sphingosine kinase (SK) homologues SKA, SKB and SKC. The
XX  polypeptides can be obtained using recombinant DNA methods, and host
XX  cells containing expression vectors including SK polynucleotides are used
XX  in a claimed method of screening for compounds that inhibit or activate
XX  human SK activity. Human SK specific antibodies, inhibitors, ligands or
XX  their analogues can be used as bioactive agents to treat inflammation or
XX  disease including viral, bacterial or fungal infections, allergic
XX  responses, mechanical injury associated with trauma, hereditary diseases,
XX  lymphoma or carcinoma, and other conditions with activate the genes of
XX  kidney, lung, heart, lymphoid or tissues of the nervous system
XX
XX  Sequence 460 AA;
XX
XX  Query Match 81.4%; Score 2463; DB 3; Length 460;
XX  Best Local Similarity 99.6%; Pred. No. 1.2e-228;
XX  Matches 458; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX  103 MEKPYAFVTVCHVKARRHRWKWAQVTFWCPPEQLCHLWQLTRMLEKLTSRPKHLVFI 162

```

```

Db      1  MEKPYAFVTVCHVKARRHRWKWAQVTFWCPPEQLCHLWQLTRMLEKLTSRPKHLVFI 60
QY      163  NPFGGKGGKRIYERKVAFLFTLASITTDIIVTEHANQAKETLYEINIDKYDGI VCVGGD 222
Db      61  NPFGGKGGKRIYERKVAFLFTLASITTDIIVTEHANQAKETLYEINIDKYDGI VCVGGD 120
QY      223  GMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGSTDVCYCYSTVGTSDAETS 282
Db      121  GMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGSTDVCYCYSTVGTSDAETS 180
QY      283  ALHIVVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDI IKDSEKKRWGLARVDFSLK 342
Db      181  ALHIVVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDI IKDSEKKRWGLARVDFSLK 240
QY      343  TFLSHHCYEGTVPFLPAQHTVGSPPDRKPCRAGFCVCRQSKQLEEEQKALYGLEAAED 402
Db      241  TFLSHHCYEGTVPFLPAQHTVGSPPDRKPCRAGFCVCRQSKQLEEEQKALYGLEAAED 300
QY      403  VEEQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFRLIR 462
Db      301  VEEQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFRLIR 360
QY      463  HTNQDQDFDTFVEVYRVKQFTSKMEDESDLKEGGKRFHICSSHPSCCCTVSN 522
Db      361  HTNQDQDFDTFVEVYRVKQFTSKMEDESDLKEGGKRFHICSSHPSCCCTVSN 420
QY      523  SWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHSHS 562
Db      421  SWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHSHS 460

RESULT 9
AAE07884
ID  AAE07884 standard; protein; 471 AA.
AC  AAE07884;
XX
XX  01-NOV-2001 (first entry)
XX  Human sphingosine kinase (SphK) protein #2.
XX
XX  Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy;
XX  anticense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
XX  cyclostatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
XX  leukaemia; vasotropic; cell proliferative disorder; vascular disease.
XX
XX  Homo sapiens.
XX
XX  WO200160990-A2.
XX
XX  23-AUG-2001.
XX
XX  14-FEB-2001; 2001WO-US004789.
XX
XX  14-FEB-2000; 2000US-0182360P.
XX  22-MAR-2000; 2000US-0191261P.
XX  (CURA-) CURAGEN CORP.
XX  (GETH ) GENENTECH INC.
XX
XX  Rastelli L;
XX
XX  WPI; 2001-514770/56.
XX  N-PSDB; AAD14426.
XX
XX  An isolated Sphingosine kinase polypeptide useful for treating a SphK-
XX  associated disorder especially cancer, restenosis or ischemia in a human.
XX
XX  Claim 1; Fig 1; 107pp; English.
XX
XX  The present invention relates to sphingosine kinase (SphK) polypeptides
XX  and nucleic acids encoding them. SphK is useful for treating a SphK-
XX  associated disorder especially cancers such as leukaemia, lymphoma,

```



```
QY 326 EKRWGLARYDFSLGKTLFSLHCYEGTVSFLPAQHTVGSPRDRKPCBAGCFVCRQSKQ 385
DB 301 EKRWGLRVYDFSLGKTLFSLHCYEGTVSFLPAQHTVGSPRDRKPCBAGCFVCRQSKQ 360
QY 386 LEEEOKALYGLEAAEDVEEVQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGL 445
DB 361 LEEEEKALYGLEAAEEVQVTCGKFLAINATNMSCACRSPGCLSPFAHLGDSGL 420
QY 446 ILIRKCSFNFRLIRHTNQDQDFTFVEVYRVKFKQFTSKMEDESDSLKGGKKRP 505
DB 421 ILIRKCSFNFRLIRHTNQDQDFTFVEVYRVKFKFTSKMEDESDSLKGGKKRP 480
QY 506 GHICSHSPSCCTVSNSSWNCDEVLHSPAEVVRVHCOLVRLFARGIEE 554
DB 481 GKICKDRPSCTSASRSWNCDEVMHSPAEVVRVHCOLVRLFARGIEE 529

RESULT 11
ADSL11054
ID ADSL11054 standard; protein; 536 AA.
XX AC ADSL11054;
XX DT 16-DEC-2004 (first entry)
XX DE Human therapeutic protein - SEQ ID 1291.
XX KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;
XX KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
XX KW aplastic anaemia; cancer; wound healing; gene therapy.
XX OS Homo sapiens.
XX FN WO2004080148-A2.
XX PD 23-SEP-2004.
XX PF 30-SEP-2003; 2003WO-US030720.
XX PR 02-OCT-2002; 2002US-0416186P.
XX PA (NUVE-) NUVELO INC.
XX FI Tang YT, Asundi V, Ren F, Zhang J, Zhang J, Wehrman T, Wang Z, Ma Y;
XX FI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX DR WPI; 2004-668857/65.
XX DR N-PSDB; ADS10370.
XX PT New polynucleotide, useful in preparing a composition for diagnosing or
XX PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
XX PT aplastic anemia or cancer for promoting wound healing.
XX PS Claim 20; SEQ ID NO 1291; 718pp; English.
XX CC The invention relates to a novel isolated polynucleotide and the encoded
XX CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
XX CC neuroprotective, antianaemic, cytostatic and vulnery activities and may
XX CC be useful in preparing a composition for diagnosing or treating
XX CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
XX CC disorders, such as aplastic anaemia or cancer, as well as for promoting
XX CC wound healing. The molecules may also be utilised during gene therapy
XX CC procedures. The current sequence is that of a human therapeutic protein
XX CC of the invention. The current sequence is not shown explicitly within the
XX CC specification but can be accessed from the WIPO web-site.
XX SQ Sequence 536 AA;

Query Match 77.7%; Score 2350; DB 8; Length 536;
Best Local Similarity 78.1%; Pred. No. 1.3e-217;
Matches 452; Conservative 2; Mismatches 9; Indels 116; Gaps 4;

QY 57 LRWR-----SPGAGAGAGACSVSPVSEIIAIVETDVHGKQGGKQWQMKPYAFT 110
```

```
DB 1 MEMGRRVFCVEMPAGRGV-FTDACSVPSVEIIAIVETDVHGKQGGKQWQMKPYAFT 59
QY 111 VHCYKRRRRHWKWAQVTFWCPPEQLCHLMQLTLEMLEKLTSPRQHLVFINPFGKGQ 170
DB 60 VHCYKRRRRHWKWAQVTFWCPPEQLCHLMQLTLEMLEKLT----- 100
QY 171 GKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKYDGIIVCGDGMFSEVLH 230
DB 101 -----ITEHANOAKETLYEINIDKYDGIIVCGDGMFSEVLH 137
QY 231 GLIGRTORSAGVDQNHPRAVLPSSLRIGIIPA----- 263
DB 138 GLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGFANDLAGDEVSVLSPVSGPEGGMV 197
QY 264 -----GSTDCVCYVTVGTSDAETSA 283
DB 198 HTFTLVLTALGCEHRSHPHFMDERRTGEHVPGENAGGLDGDSTDCVCYVTVGTSDAETSA 257
QY 284 LHIIVGDSLAMDVSSVHHNSTLLRYSVLLGYGYGDI IKDSEKKRWGLARYDFSLGKT 343
DB 258 LHIIVGDSLAMDVSSVHHNSTLLRYSVLLGYGYGDI IKDSEKKRWGLARYDFSLGKT 317
QY 344 FLSHHCYEGTVSFLPAQHTVGSPRDRKPCBAGCFVCRQSKQLEEOKKALYGLEAAEDV 403
DB 318 FLSHHCYEGTVSFLPAQHTVGSPRDRKPCBAGCFVCRQSKQLEEOKKALYGLEAAEDV 377
QY 404 EEMOVVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGLIIRKCSFNFRLIRH 463
DB 378 EEMOVVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGLIIRKCSFNFRLIRH 437
QY 464 TNQODQDFTFVEVYRVKFKQFTSKMEDESDSLKGGKKRFGHICSHSPSCCTVSNSS 523
DB 438 TNQODQDFTFVEVYRVKFKQFTSKMEDESDSLKGGKKRFGHICSHSPSCCTVSNSS 497
QY 524 WNCDEVLHSPAEVVRVHCOLVRLFARGIEENPKPDSHS 562
DB 498 WNCDEVLHSPAEVVRVHCOLVRLFARGIEENPKPDSHS 536

RESULT 12
ABG13541
ID ABG13541 standard; protein; 746 AA.
XX AC ABG13541;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #13532.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX FN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX FI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX DR N-PSDB; AAS77728.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
```

PT responsible for genetic disorders or other traits and to assess
XX biodiversity.

PS Claim 20; SEQ ID NO 43900; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 746 AA;

Query Match 73.2%; Score 2215; DB 4; Length 746;
Best Local Similarity 68.7%; Pred. No. 2.3e-204;
Matches 444; Conservative 12; Mismatches 44; Indels 146; Gaps 9;

```
Qy 6 GPAPLGVRAPAWRTSPAEMGATGAEPLOSILWVKQOR---CAVSLPARAL----- 56
Db 124 GPAPTL---PHSSDVPPEARAGDRPSSQLTLHAGLSQRACCGSLVSPVHCLFLVSCP 180
Qy 57 -----LRWWR-----SPGPGAGPGADACSPVSEIIAIVEETDVH 91
Db 181 ELSVIAGREADAGFRPDSMEWGRVFCVEMPAGERGV-FTDACSVPVSEIIAIVEETDVH 239
Qy 92 GHQSGGKQWQKXPYAFTHVCVKRRHRKWAQVTFWCPEQLCHLWLTQTLREMLEKL 151
Db 240 GHQSGGKQWQKXPYAFTHVCVKRRHRKWAQVTFWCPEQLCHLWLTQTLREMLEKL 299
Qy 152 TSPKPHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINID 211
Db 300 -----ITEHANOAKETLYEINID 317
Qy 212 KYDGIVCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA----- 263
Db 318 KYDGIVCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPAETHVGPGEN 377
Qy 264 -----GSTDCVCYSTVGTSDAETSALHIV----- 288
Db 378 AGGLDGDSTDCVCYSTVGTSDAETSALHIVGCCPEARKPPASRHTACSGHGQLCLGQOL 437
Qy 289 -----GDSLAMDVSSVHNSLTLLRYSVSLIGYGYGDIIDKSEKRW 330
Db 438 WNACLCBASLQSRMQSPGDSLAMDVSSVHNSLTLLRYSVSLIGYGYGDIIDKSEKRW 497
Qy 331 LGLARYDFSGUKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQOLEEQ 390
Db 498 LGLARYDFSGUKTFLSHHCYEGTGSFLPAQHTVGSPRNRKPCRCAGCFVCRQSKQOLEEQ 557
Qy 391 KKALYGLEAEADVEEQVVCVKFLAINATNMSCACRSPRGLSPAHLGDSGLILIRK 450
Db 558 KKALYGLEAEADVEEQVVCVKFLAINATNMSCACRSPRGLSPAHLGDSGLILIRK 617
Qy 451 CSRNFRLRLIRHTNQODQDFTFVEYVRVKKFOFTSKHMEDESDLKEGKKRPFHICS 510
Db 618 CSRNFRLRLIRHTNQODQDFTFVEYVRVKKFLFTSKHMEDESDLKEGKKRPFHICS 677
```

Qy 511 SHPSCCCTVSNSSWNCDEVLHSPHSPAJEVRVHCOLVRLFARGIENP 556
Db 678 SHPSCCCTVSNSSWNCDEVLHSPAJEVR-----AVVMARGCRKVP 718

RESULT 13
ABRS6301

ID ABR56301 standard; protein; 416 AA.

XX ABR56301;

DT 20-NOV-2003 (first entry)

DE Human Sphingosine kinase 4.

KW Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;

KW platelet transfusion; platelet stabiliser.

OS Homo sapiens.

PN WO2003031627-A1.

XX 17-APR-2003.

XX 28-SEP-2001; 2001WO-JP008537.

XX 28-SEP-2001; 2001WO-JP008537.

XX (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.

PA (CHBI-) CHEM BIOLOGY INST.

PI Igarashi Y, Kihara A;

DR WPI; 2003-354917/33.

DR N-PSDB; ACC70838.

XX Platelet derived polypeptides with sphingosine kinase activity for
PT treatment of sphingosine related disorders.

XX Claim 2; Page 29-30; 39pp; Japanese.

XX The present sequence is the protein sequence for human sphingosine kinase
CC 4 (SPHK4). The kinase can be used for the diagnosis and treatment of
CC sphingosine related disorders. The kinase can also be potentially used
CC for controlling toxicity of platelet transfusion and as a platelet
CC stabiliser

XX Sequence 416 AA;

Query Match 73.1%; Score 2210; DB 6; Length 416;
Best Local Similarity 100.0%; Pred. No. 3e-204;
Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 147 MLEKLSRPRKHLVFNPFPGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLY 206
Db 1 MLEKLSRPRKHLVFNPFPGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLY 60
Qy 207 EINDIKYDGIVCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPAGST 266
Db 61 EINDIKYDGIVCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPAGST 120
Qy 267 DCVCYSTVGTSDAETSALHIVGDSLAMDVSSVHNSLTLLRYSVSLIGYGYGDIIDKSE 326
Db 121 DCVCYSTVGTSDAETSALHIVGDSLAMDVSSVHNSLTLLRYSVSLIGYGYGDIIDKSE 180
Qy 327 KKWGLGLARYDFSGUKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQOL 386
Db 181 KKWGLGLARYDFSGUKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQOL 240
Qy 387 EEOKKALYGLEAEADVEEQVVCVKFLAINATNMSCACRSPRGLSPAHLGDSGLI 446
Db 241 EEOKKALYGLEAEADVEEQVVCVKFLAINATNMSCACRSPRGLSPAHLGDSGLI 300
```


QY 447 LIRKCSRFNLFRLIRHTNQDQDFTFVEVVRVKKFOFTSKHMEDESDLKEGKKEPFG 506
|||||
Db 301 LIRKCSRFNLFRLIRHTNQDQDFTFVEVVRVKKFOFTSKHMEDESDLKEGKKEPFG 360
|||||
QY 507 HICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIENPKPDGSHS 562
|||||
Db 361 HICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIENPKPDGSHS 416
|||||

RESULT 14

ABGI3543
ID ABGI3543 standard; protein; 727 AA.

XX AC ABGI3543;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #13534.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS77730.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX PS Claim 20; SEQ ID NO 43902; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 727 AA;

Query Match 56.7%; Score 1714.5; DB 4; Length 727;
Best Local Similarity 63.7%; Pred. No. 5.2e-156;

Matches 345; Conservative 0; Mismatches 0; Indels 197; Gaps 3;
QY 195 TEHANQAKETLYEINIDKYDG----- 215
|||||
Db 1 TEHANQAKETLYEINIDKYDGYTAHSRMPCPNRNHSAGKAAAPVQGVGYGSGQRAFPWEP 60
|||||
QY 216 -----IVCVCGDCGMFSEVLHGL 232
|||||
Db 61 GPKOMPAQGVCRPLPLKLEASGLLRSEGRCTCRSPDLRCSSCSIVCVCGDCGMFSEVLHGL 120
|||||
QY 233 IGRTORAGVDQNHPRAVLVPSSLRIGIIPA----- 263
|||||
Db 121 IGRTORAGVDQNHPRAVLVPSSLRIGIIPAGFANDLAGDEVSVLSGPVSGEGGMVHT 180
|||||
QY 264 -----GSTDCVCYSTVGTSDAETSALH 285
|||||
Db 181 FTLVTALGCEHRSHPHFMEDEDRRTGEHVHVGPNAGGLDGSTDCVCYSTVGTSDAETSALH 240
|||||
QY 286 IVV-----GDSLAMDVS 297
|||||
Db 241 IVVCCPEARKPPASRHTACSGHGQLCLGQLWNACLCEASRLQSRMQSPGDSLAMDVS 300
|||||
QY 298 SVHNSHTLLRYSVSLLGVGYGDIIDKSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFL 357
|||||
Db 301 SVHNSHTLLRYSVSLLGVGYGDIIDKSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFL 360
|||||
QY 358 PAQTVGSPDRKPCRAGCFVCRQSKQLEBEQKALYGLEAAEDVEWQVVCCKFLAIN 417
|||||
Db 361 PAQTVGSPDRKPCRAGCFVCRQSKQLEBEQKALYGLEAAEDVEWQVVCCKFLAIN 420
|||||
QY 418 ATNMSCACRRSPRGLSPAHLGDGSSDILIRKCSRNFRLIRHTNQDQDFTFVEV 477
|||||
Db 421 ATNMSCACRRSPRGLSPAHLGDGSSDILIRKCSRNFRLIRHTNQDQDFTFVEV 480
|||||
QY 478 YRVKKFOFTSKHMEDESDLKEGKKEPFGHICSSHPSCCTVSNSSWNCDEVLHSPAIE 537
|||||
Db 481 YRVKKFOFTSKHMEDESDLKEGKKEPFGHICSSHPSCCTVSNSSWNCDEVLHSPAIE 540
|||||
QY 538 VR 539
||
Db 541 VR 542
|||||
RESULT 15
ABB07854
ID ABB07854 standard; protein; 326 AA.
XX AC ABB07854;
XX DT 03-JUL-2002 (first entry)
XX DE Human sphingosine kinase-like protein.
XX KW Human sphingosine kinase-like protein; intracellular signalling;
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease.
XX OS Homo sapiens.
XX PN WO200228906-A2.
XX PD 11-APR-2002.
XX PF 05-OCT-2001; 2001WO-EP011516.
XX PR 06-OCT-2000; 2000US-0238005P.
XX PR 23-AUG-2001; 2001US-0314113P.
XX PA (FARB) BAYER AG.
XX PI Kossida S, Encinas J;
XX DR WPI; 2002-340094/37.

DR N-PSDB; ABL40822.
XX New reagent for modulating the activity of sphingosine kinase-like
PT protein polypeptide or polynucleotide and treating cancer, asthma,
PT allergy, an autoimmune disease, or a central or peripheral nervous system
XX disorder.
XX
XX
PS Claim 25; Fig 2; 120pp; English.
XX
XX The invention relates to a human sphingosine kinase-like protein. The
CC polypeptide can be expressed by standard recombinant methodology. The
CC sphingosine kinase-like protein and gene can be used to regulate
CC intracellular signalling and consequently cell proliferation and
CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
CC peripheral nervous system disorders (e.g. Parkinson's disease). The
CC present sequence represents the human sphingosine kinase-like protein
XX
SQ Sequence 326 AA;
Query Match 54.2%; Score 1640.5; DB 5; Length 326;
Best Local Similarity 96.3%; Pred. No. 2.3e-149;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
Qy 155 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 203
Db 1 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFVYVYVEVITEHANOAKE 60
Qy 204 TLYEINIDKYDGIYVCGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 263
Db 61 TLYEINIDKYDGIYVCGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 120
Qy 264 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLRYSVSLGCGFYGDIIK 323
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLRYSVSLGCGFYGDIIK 180
Qy 324 DSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSK 383
Db 181 DSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSK 240
Qy 384 QQLBEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSS 443
Db 241 QQLBEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
Qy 444 DLILIRKCSRFNLFRLIRHTNQDDQ 469
Db 301 DLILIRKCSRFNLFRLIRHTNQDDQ 326

Search completed: September 3, 2005, 04:37:42
Job time : 116.766 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:08:02 ; Search time 111.217 Seconds
(without alignments)
1990.064 Million cell updates/sec

Title: US-10-631-958-11

Perfect score: 3025

Sequence: 1 HEAANGAPLGVAPPAPWRT.....QLVRLFARGIEENKPDHS 562

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3025	100.0	562	10	US-09-969-896-11
2	3025	100.0	562	16	US-10-631-958-11
3	2888	95.5	537	10	US-09-969-896-10
4	2888	95.5	537	15	US-10-262-511-40
5	2888	95.5	537	16	US-10-631-958-10
6	2888	95.5	537	16	US-10-618-941-121
7	2880	95.2	537	14	US-10-315-597A-2
8	2456.5	81.2	471	9	US-09-784-810A-6
9	2456.5	81.2	471	18	US-10-876-281-6
10	1640.5	54.2	326	9	US-09-784-810A-11
11	1640.5	54.2	326	10	US-09-969-896-2

12	1640.5	54.2	326	16	US-10-631-958-2	Sequence 2, Appli
13	1640.5	54.2	326	18	US-10-876-281-11	Sequence 11, Appl
14	986	32.6	182	9	US-09-784-810A-29	Sequence 29, Appl
15	986	32.6	182	18	US-10-876-281-29	Sequence 29, Appl
16	647	21.4	144	9	US-09-784-810A-8	Sequence 8, Appli
17	647	21.4	144	18	US-10-876-281-8	Sequence 8, Appli
18	611.5	20.2	588	16	US-10-425-115-361076	Sequence 361076,
19	597.5	19.8	575	16	US-10-425-115-270874	Sequence 270874,
20	480.5	15.9	613	16	US-10-437-963-195284	Sequence 195284,
21	404.5	13.4	484	16	US-10-437-963-133133	Sequence 133133,
22	389	12.9	487	15	US-10-424-599-190238	Sequence 190238,
23	362.5	12.0	760	16	US-10-425-115-213665	Sequence 213665,
24	357.5	11.8	490	16	US-10-425-115-249553	Sequence 249553,
25	344.5	11.4	381	16	US-10-425-115-249555	Sequence 249555,
26	341	11.3	520	15	US-10-424-599-202261	Sequence 202261,
27	335	11.1	79	9	US-09-784-810A-10	Sequence 10, Appl
28	335	11.1	79	18	US-10-876-281-10	Sequence 10, Appl
29	334	11.0	245	15	US-10-425-114-53957	Sequence 53957, A
30	331.5	11.0	446	16	US-10-437-963-197631	Sequence 197631,
31	330.5	10.9	757	16	US-10-437-963-180379	Sequence 180379,
32	326.5	10.8	470	16	US-10-425-115-282512	Sequence 282512,
33	326.5	10.8	521	15	US-10-425-114-68423	Sequence 68423, A
34	313	10.3	552	16	US-10-437-963-161184	Sequence 161184,
35	313	10.3	677	16	US-10-425-115-264087	Sequence 264087,
36	312.5	10.3	373	16	US-10-767-701-44927	Sequence 44927, A
37	310.5	10.3	641	15	US-10-348-052-28	Sequence 28, Appl
38	310.5	10.3	641	16	US-10-622-011-28	Sequence 28, Appl
39	306	10.1	490	14	US-10-053-510-19	Sequence 19, Appl
40	306	10.1	490	15	US-10-348-052-19	Sequence 19, Appl
41	306	10.1	490	16	US-10-622-011-19	Sequence 19, Appl
42	301	10.0	524	14	US-10-053-510-20	Sequence 20, Appl
43	301	10.0	524	15	US-10-348-052-20	Sequence 20, Appl
44	301	10.0	524	16	US-10-622-011-20	Sequence 20, Appl
45	301	10.0	907	15	US-10-348-052-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1

US-09-969-896-11
; Sequence 11, Application US/09969896
; Publication No. US2003012553A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-896-11

Query Match	100.0%	Score 3025;	DB 10;	Length 562;
Best Local Similarity	100.0%	Pred. No. 5.2e-281;		
Matches	562;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Oy	1	HEAANGAPLGVAPPAPWRTSPAEMGATGAAPLQSVLVWVKQRCVLSLEPAPALLRW 60		
Db	1	HEAANGAPLGVAPPAPWRTSPAEMGATGAAPLQSVLVWVKQRCVLSLEPAPALLRW 60		
Oy	61	RSPGPGAGACADACSPVSEIIAIVETDVHGKQSGKQMKPKYFTVHCVRARRH 120		
Db	61	RSPGPGAGACADACSPVSEIIAIVETDVHGKQSGKQMKPKYFTVHCVRARRH 120		

QY	121	RKWAQVTFWCPEBQI	CHLWLOTREMLEKLT	SRPKHLVFINP	PGKGQGGKRIYERKVA	180
Db	121	RKWAQVTFWCPEEQI	CHLWLOTREMLEKLT	SRPKHLVFINP	PGKGQGGKRIYERKVA	180
QY	181	PUFTLASITTDI	IIVTEHANQAKETLYE	INIDKYDGCIVCVGGDGMFSEVHLHGLIGRTQSA	240	
Db	181	PUFTLASITTDI	IIVTEHANQAKETLYE	INIDKYDGCIVCVGGDGMFSEVHLHGLIGRTQSA	240	
QY	241	GVQDQNHPRAVLPV	SSLRIGIIPAGSTDCV	CVYVGTSDAETSAHLIIVGDSLAMDVSSVH	300	
Db	241	GVQDQNHPRAVLPV	SSLRIGIIPAGSTDCV	CVYVGTSDAETSAHLIIVGDSLAMDVSSVH	300	
QY	301	HNSTLLRYSVSL	LGFGFYCGDI	IKDSEKKRWLGLARYDF	SGLKTFLSHHCYEGTVSFLPAQ	360
Db	301	HNSTLLRYSVSL	LGFGFYCGDI	IKDSEKKRWLGLARYDF	SGLKTFLSHHCYEGTVSFLPAQ	360
QY	361	HTVGSPRDRKPC	RACGFCVCRQSKO	OLEBEOKALYGLEAAEDVEEQVVCCKFLAINATN	420	
Db	361	HTVGSPRDRKPC	RACGFCVCRQSKO	OLEBEOKALYGLEAAEDVEEQVVCCKFLAINATN	420	
QY	421	MSCACRRSPRGLS	PAAHLDGSSDIL	ILIRKCSRNFRLIRHTNQDQDFDFTFVEVYRV	480	
Db	421	MSCACRRSPRGLS	PAAHLDGSSDIL	ILIRKCSRNFRLIRHTNQDQDFDFTFVEVYRV	480	
QY	481	KKQFTSKHMEDESD	LKEGKKRFGHIC	SSHPSCCTVNSNWCNCDGEVLHSPAIEVRV	540	
Db	481	KKQFTSKHMEDESD	LKEGKKRFGHIC	SSHPSCCTVNSNWCNCDGEVLHSPAIEVRV	540	
QY	541	HCQVRLFARGIEEN	PKPDSSH	562		
Db	541	HCQVRLFARGIEEN	PKPDSSH	562		

```

RESULT 2
US-10-631-958-11
; Sequence 11, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631.958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-958-11

```

Db	121	RWKWAQVTFWCPBEQOLCHLWLQILREMLEKLTSRPKHLLVFINPPGKGQGGKRIYERKVA	180
Qy	181	PLFTLASITTDIIIVTEHANOAKETLYBEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSA	240
Db	181	PLFTLASITTDIIIVTEHANOAKETLYBEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSA	240
Qy	241	GVDQNHPRAVLVPSSLRIGIIPAGSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSVVH	300
Db	241	GVDQNHPRAVLVPSSLRIGIIPAGSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSVVH	300
Qy	301	HNSLTLLRYSVSLLCYGIFYGDIIDKSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQ	360
Db	301	HNSLTLLRYSVSLLCYGIFYGDIIDKSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQ	360
Qy	361	HTVGSPRDRKPCRAGCFVCRQSKQOLEEBEOKKALYGLEAAEDVEEMOVVCGKFLAINATN	420
Db	361	HTVGSPRDRKPCRAGCFVCRQSKQOLEEBEOKKALYGLEAAEDVEEMOVVCGKFLAINATN	420
Qy	421	MSCACRSRGLSPAHLGPGSSDLILIRKCSRFNFRFLIRHTNODQDFDFVEVYVR	480
Db	421	MSCACRSRGLSPAHLGPGSSDLILIRKCSRFNFRFLIRHTNODQDFDFTFVEVYVR	480
Qy	481	KKFOFTSKHWEDESDLKEGKKRFGHCSSHPSCCCTVSNSSNWCNCDGEVLHSPAIEVRV	540
Db	481	KKFOFTSKHWEDESDLKEGKKRFGHCSSHPSCCCTVSNSSNWCNCDGEVLHSPAIEVRV	540
Qy	541	HCQLVRLFARGIENPKPDSHS	562
Db	541	HCQLVRLFARGIENPKPDSHS	562

```

RESULT 3
US-09-969-896-10
; Sequence 10, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: fastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-896-10

```

	Query Match	95.5%	Score 2888;	DB 10;	Length 537;
	Best Local Similarity	100.0%;	Pred. No. 6.9e-268;		
	Matches 537; Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	26	MGATGAAPQLQSVLVWKQORCAVSLEPARALLRWRSPGCGAGAPGADACSVVPSEIIAV	85		
Dd	1	MGATGAAPLQSLVWLKQORCAVSLEPARALLRWRSPGCGAGAPGADACSVVPSEIIAV	60		
Qy	86	EETDVHGKGSGKWQKMEKPYAFTVHCVKRARHRHWKAQVTFWCPEEQICHUWLOTLR	145		
Dd	61	EETDVHGKGSGKWQKMEKPYAFTVHCVKRARHRHWKAQVTFWCPEEQICHUWLOTLR	120		
Qy	146	EMLEKLTSRPKHLIAVFNPFGKGQGKKRIYERKVAPLFTLIASITTDIIVTEHANQAETL	205		
Dd	121	EMLEKLTSRPKHLIAVFNPFGKGQGKKRIYERKVAPLFTLIASITTDIIVTEHANQAETL	180		
Qy	206	YEINIDKYDGIVCVGGDGMSFVHLHIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS	265		
Dd	181	YEINIDKYDGIVCVGGDGMSFVHLHIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS	240		

```

Qy 266 TDCVYSTVCTSDAETSAHIVVGDSLANDVSSVHHNSTLLRYSVSLGCGFYGDIKDS 325
Db 241 TDCVYSTVGTSGDAETSAHIVVGDSLANDVSSVHHNSTLLRYSVSLGCGFYGDIKDS 300
Qy 326 EKKRWGLGARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 385
Db 301 EKKRWGLGARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 360
Qy 386 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGL 445
Db 361 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGL 420
Qy 446 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKEGGKRF 505
Db 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKEGGKRF 480
Qy 506 GHICSSHPSCCTVSNSSWNCDEVLHSPATEVRVHCOLVRLFARGIEENPKPDHS 562
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPATEVRVHCOLVRLFARGIEENPKPDHS 537

RESULT 4
US-10-262-511-40
; Sequence 40, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; PRIOR FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056

```

```

; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 40
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-40

Query Match 95.5%; Score 2888; DB 15; Length 537;
Best Local Similarity 100.0%; Pred. No. 6,9e-268;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAPLQSVLWVKQRCVSLPARALLRWRSFPGAGAPGADACSVPVSEIIAV 85
Db 1 MGATGAAPLQSVLWVKQRCVSLPARALLRWRSFPGAGAPGADACSVPVSEIIAV 60

Qy 86 EETDVHGKHOCSGKQWMEKPYAFTVHCVRARHRWKQAQVTFWCPEEQLCMLWLTLR 145
Db 61 EETDVHGKHOCSGKQWMEKPYAFTVHCVRARHRWKQAQVTFWCPEEQLCMLWLTLR 120

Qy 146 EMLEKLSRPAKHLVFINPFGKGQGRKRIYERKVAFLFTLASITTDIIIVTSHANOAKETL 205
Db 121 EMLEKLSRPAKHLVFINPFGKGQGRKRIYERKVAFLFTLASITTDIIIVTSHANOAKETL 180

Qy 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRVAVLPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRVAVLPSSLRIGIIPAGS 240

Qy 266 TDCVYSTVGTSDAETSAHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIKDS 325
Db 241 TDCVYSTVGTSDAETSAHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIKDS 300

Qy 326 EKKRWGLGARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 385
Db 301 EKKRWGLGARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 360

Qy 386 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGL 445
Db 361 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGL 420

Qy 446 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKEGGKRF 505
Db 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKEGGKRF 480

Qy 506 GHICSSHPSCCTVSNSSWNCDEVLHSPATEVRVHCOLVRLFARGIEENPKPDHS 562
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPATEVRVHCOLVRLFARGIEENPKPDHS 537

RESULT 5
US-10-631-958-10
; Sequence 10, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113

```

		PRIORITY FILING DATE: 2001-08-23		Best Local Similarity 100.0%; Pred. No. 6.9e-268;		Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
		NUMBER OF SEQ ID NOS: 16					
		SOFTWARE: FastSeq for Windows Version 4.0					
		SEQ ID NO 10					
		LENGTH: 537					
		TYPE: PRT					
		ORGANISM: Homo sapiens					
		US-10-631-958-10					
		Query Match 95.5%; Score 2888; DB 16; Length 537;					
		Best Local Similarity 100.0%; Pred. No. 6.9e-268;					
		Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	26	MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	85				
Db	1	MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	60				
Qy	86	EETDVHGKQSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	145				
Db	61	EETDVHGKQSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	120				
Qy	146	EMLEKLTSPKHLVFINPFGKGQGRKIYERKVAFLFTLASITTDIIIVTEHANQAKETL	205				
Db	121	EMLEKLTSPKHLVFINPFGKGQGRKIYERKVAFLFTLASITTDIIIVTEHANQAKETL	180				
Qy	206	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS	265				
Db	181	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS	240				
Qy	266	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	325				
Db	241	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	300				
Qy	326	EKKRWGLIARYDFSGIKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCACFCVCRQSKQ	385				
Db	301	EKKRWGLIARYDFSGIKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCACFCVCRQSKQ	360				
Qy	386	LEEQKKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	445				
Db	361	LEEQKKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	420				
Qy	446	ILIRKCSRFNPLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	505				
Db	421	ILIRKCSRFNPLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	480				
Qy	506	GHICSSHPSCCTVSNSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDHS	562				
Db	481	GHICSSHPSCCTVSNSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDHS	537				
		RESULT 7					
		US-10-315-597A-2					
		Sequence 2, Application US/10315597A					
		Publication No. US20030162206A1					
		GENERAL INFORMATION:					
		APPLICANT: Sugiyura, Masako					
		APPLICANT: Kono, Keita					
		TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It					
		FILE REFERENCE: 02658CIP/HG					
		CURRENT APPLICATION NUMBER: US/10/315,597A					
		CURRENT FILING DATE: 2002-12-10					
		PRIOR APPLICATION NUMBER: JP 2000-178039					
		PRIOR FILING DATE: 2000-06-14					
		NUMBER OF SEQ ID NOS: 4					
		SOFTWARE: PatentIn Ver. 2.0					
		SEQ ID NO 2					
		LENGTH: 537					
		TYPE: PRT					
		ORGANISM: Homo sapiens					
		US-10-315-597A-2					
		Query Match 95.2%; Score 2880; DB 14; Length 537;					
		Best Local Similarity 99.6%; Pred. No. 4.1e-267;					
		Matches 535; Conservative 1; Mismatches 1; Indels 0; Gaps 0;					
Qy	26	MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	85				
Db	1	MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	60				
Qy	86	EETDVHGKQSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	145				
Db	61	EETDVHGKQSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	120				
		Query Match 95.5%; Score 2888; DB 16; Length 537;					
		Best Local Similarity 100.0%; Pred. No. 6.9e-268;					
		Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	26	MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	85				
Db	1	MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	60				
Qy	86	EETDVHGKQSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	145				
Db	61	EETDVHGKQSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	120				
Qy	146	EMLEKLTSPKHLVFINPFGKGQGRKIYERKVAFLFTLASITTDIIIVTEHANQAKETL	205				
Db	121	EMLEKLTSPKHLVFINPFGKGQGRKIYERKVAFLFTLASITTDIIIVTEHANQAKETL	180				
Qy	206	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS	265				
Db	181	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS	240				
Qy	266	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	325				
Db	241	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	300				
Qy	326	EKKRWGLIARYDFSGIKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCACFCVCRQSKQ	385				
Db	301	EKKRWGLIARYDFSGIKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCACFCVCRQSKQ	360				
Qy	386	LEEQKKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	445				
Db	361	LEEQKKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	420				
Qy	446	ILIRKCSRFNPLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	505				
Db	421	ILIRKCSRFNPLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	480				
Qy	506	GHICSSHPSCCTVSNSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDHS	562				
Db	481	GHICSSHPSCCTVSNSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDHS	537				
		RESULT 6					
		US-10-618-941-121					
		Sequence 121, Application US/10618941					
		Publication No. US2004019792A1					
		GENERAL INFORMATION:					
		APPLICANT: WHYTE, DAVID					
		APPLICANT: MANNING, GERARD					
		APPLICANT: CAENEPEEL, SEAN					
		TITLE OF INVENTION: NOVEL KINASES					
		FILE REFERENCE: 034536-0321					
		CURRENT APPLICATION NUMBER: US/10/618,941					
		CURRENT FILING DATE: 2003-07-15					
		PRIOR APPLICATION NUMBER: 60/395,632					
		PRIOR FILING DATE: 2002-07-15					
		NUMBER OF SEQ ID NOS: 143					
		SOFTWARE: PatentIn version 3.2					
		SEQ ID NO 121					
		LENGTH: 537					
		TYPE: PRT					
		ORGANISM: Homo sapiens					
		US-10-618-941-121					
		Query Match 95.5%; Score 2888; DB 16; Length 537;					

RESULT 14

Query Match 32.6%; Score 986; DB 18; Length 182;

Best Local Similarity 100.0%; Pred. No. 6.4e-86;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	67	AGAPGADACSVPVSEIIIAVEETDVHGKHQSGKWQMEKPYAFTVHCVKRARRHRWKWAQ	126
Db	1	AGAPGADACSVPVSEIIIAVEETDVHGKHQSGKWQMEKPYAFTVHCVKRARRHRWKWAQ	60
Qy	127	VTFCPEEQCHLWLQTLREMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAPLFTLA	186
Db	61	VTFCPEEQCHLWLQTLREMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAPLFTLA	120
Qy	187	SITTDIIIVTEHANOAKETLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNH	246
Db	121	SITTDIIIVTEHANOAKETLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNH	180
Qy	247	PR 248	
Db	181	PR 182	

Search completed: September 3, 2005, 04:32:36
Job time : 112.217 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:08:02; Search time 106.269 Seconds
(without alignments)
1990.064 Million cell updates/sec

Title: US-10-631-958-10
Perfect score: 2888
Sequence: 1 MGATGAAPLQSLVWVKQQR.....QLVRLFARGIENPKPDHS 537

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2888	100.0	537	10	US-09-969-896-10
2	2888	100.0	537	15	US-10-262-511-40
3	2888	100.0	537	16	US-10-631-958-10
4	2888	100.0	537	16	US-10-618-941-121
5	2888	100.0	562	10	US-09-969-896-11
6	2888	100.0	562	16	US-10-631-958-11
7	2880	99.7	537	14	US-10-315-597A-2
8	2456.5	85.1	471	9	US-09-784-810A-6
9	2456.5	85.1	471	18	US-10-876-281-6
10	1640.5	56.8	326	9	US-09-784-810A-11
11	1640.5	56.8	326	10	US-09-969-896-2

12	1640.5	56.8	326	16	US-10-631-958-2	Sequence 2, Appl
13	1640.5	56.8	326	18	US-10-876-281-11	Sequence 11, Appl
14	986	34.1	182	9	US-09-784-810A-29	Sequence 29, Appl
15	986	34.1	182	18	US-10-876-281-29	Sequence 28, Appl
16	647	22.4	144	9	US-09-784-810A-8	Sequence 8, Appl
17	647	22.4	144	18	US-10-876-281-8	Sequence 8, Appl
18	611.5	21.2	588	16	US-10-425-115-361076	Sequence 361076,
19	597.5	20.7	575	16	US-10-425-115-270874	Sequence 270874,
20	480.5	16.6	613	16	US-10-437-963-195284	Sequence 195284,
21	399.5	13.8	484	16	US-10-437-963-133133	Sequence 133133,
22	389	13.5	487	15	US-10-424-599-190238	Sequence 190238,
23	362.5	12.6	760	16	US-10-425-115-213665	Sequence 213665,
24	357.5	12.4	490	16	US-10-425-115-249553	Sequence 249553,
25	344.5	11.9	381	16	US-10-425-115-249555	Sequence 249555,
26	341	11.8	520	15	US-10-424-599-203261	Sequence 203261,
27	335	11.6	79	9	US-09-784-810A-10	Sequence 10, Appl
28	335	11.6	79	18	US-10-876-281-10	Sequence 10, Appl
29	334	11.6	245	15	US-10-425-114-53957	Sequence 53957, A
30	331.5	11.5	446	16	US-10-437-963-197631	Sequence 197631,
31	330.5	11.4	757	16	US-10-437-963-180379	Sequence 180379,
32	326.5	11.3	470	16	US-10-425-115-282512	Sequence 282512,
33	326.5	11.3	521	15	US-10-425-114-68423	Sequence 68423, A
34	313	10.8	677	16	US-10-425-115-264087	Sequence 264087,
35	312.5	10.8	373	16	US-10-767-701-44927	Sequence 44927, A
36	312.5	10.8	552	16	US-10-437-963-161184	Sequence 161184,
37	310.5	10.8	641	15	US-10-348-052-28	Sequence 28, Appl
38	310.5	10.8	641	16	US-10-622-011-28	Sequence 28, Appl
39	306	10.6	490	14	US-10-053-510-19	Sequence 19, Appl
40	306	10.6	490	15	US-10-348-052-19	Sequence 19, Appl
41	306	10.6	490	16	US-10-622-011-19	Sequence 19, Appl
42	301	10.4	524	14	US-10-053-510-20	Sequence 20, Appl
43	301	10.4	524	15	US-10-348-052-20	Sequence 20, Appl
44	301	10.4	524	16	US-10-622-011-20	Sequence 20, Appl
45	301	10.4	907	15	US-10-348-052-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-09-969-896-10
; Sequence 10, Application US/0969896
; Publication No. US2003012553A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974, 00594
; CURRENT APPLICATION NUMBER: US/09/969, 896
; CURRENT FILING DATE: 2001-10-04
; PRIOR FILING DATE: 2001-10-06
; PRIOR FILING DATE: 2000-10-06
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-896-10

Query Match	100.0%	Score 2888;	DB 10;	Length 537;
Best Local Similarity	100.0%	Pred. No. 3e-278;		
Matches 537;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MGATGAAPLQSLVWVKQQRCAVSLERALLRWRSPGCGAGAPGADACSVPVSEIAV	60	
Db	1	MGATGAAPLQSLVWVKQQRCAVSLERALLRWRSPGCGAGAPGADACSVPVSEIAV	60	
Qy	61	EETDVHGKQSGKWKQKXPYAFVHCVRKRHRKWAQVTFWCPEEQQLCHLWLTLR	120	
Db	61	EETDVHGKQSGKWKQKXPYAFVHCVRKRHRKWAQVTFWCPEEQQLCHLWLTLR	120	

```
Qy 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Db 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRVAVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRVAVLPSSLRIGIIPAGS 240
Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIIKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIIKDS 300
Qy 301 EKKRWGLGARYDFSLGKTFLSHHCVEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360
Db 301 EKKRWGLGARYDFSLGKTFLSHHCVEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360
Qy 361 LEEQKALYGLEAAEDVEEQVVGCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 361 LEEQKALYGLEAAEDVEEQVVGCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Qy 421 ILIRKCSRPNFLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGKKRF 480
Db 421 ILIRKCSRPNFLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGKKRF 480
Qy 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
```

RESULT 2

```
US-10-262-511-40
; Sequence 40, Application US/10262511
; Publication No. US2004038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elna
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; PRIORITY FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
```

```
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Curaseq1ist version 0.1
; SEQ ID NO 40
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-262-511-40

Query Match 100.0%; Score 2888; DB 15; Length 537;
Best Local Similarity 100.0%; Pred. No. 3e-278;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAEPLOSVLWKQQRCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60
Db 1 MGATGAAEPLOSVLWKQQRCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60
Qy 61 BETDVHGKHGSGKWKQWKPEYAFTHVCVKRRRRHWKQAQVFWCPPEOLCHLWLQTLR 120
Db 61 BETDVHGKHGSGKWKQWKPEYAFTHVCVKRRRRHWKQAQVFWCPPEOLCHLWLQTLR 120
Qy 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Db 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRVAVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRVAVLPSSLRIGIIPAGS 240
Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIIKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIIKDS 300
Qy 301 EKKRWGLGARYDFSLGKTFLSHHCVEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360
Db 301 EKKRWGLGARYDFSLGKTFLSHHCVEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360
Qy 361 LEEQKALYGLEAAEDVEEQVVGCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 361 LEEQKALYGLEAAEDVEEQVVGCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Qy 421 ILIRKCSRPNFLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGKKRF 480
Db 421 ILIRKCSRPNFLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGKKRF 480
Qy 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537

RESULT 3
US-10-631-958-10
; Sequence 10, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
```

```
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 537
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-631-958-10

Query Match      100.0%; Score 2888; DB 16; Length 537;
Best Local Similarity 100.0%; Pred. No. 3e-278;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLQSVLWVKQRCVSLPAPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60
Db 1 MGATGAAPLQSVLWVKQRCVSLPAPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60

Qy 61 EETDVHGKHQSGKQWKMEKPYAFTVHCVRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 120
Db 61 EETDVHGKHQSGKQWKMEKPYAFTVHCVRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 120

Qy 121 EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Db 121 EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180

Qy 181 YEINIDKYDGIWCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGIWCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGS 240

Qy 241 TDCVCYSTVGTSDAETSALHIVWGDLSLMDVSSVHHNSTLLRYSLVSLGFGYFGDIKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVWGDLSLMDVSSVHHNSTLLRYSLVSLGFGYFGDIKDS 300

Qy 301 EKKRWGLIARYDFSLGKTFSLSHCYEGTVSFLPAQHTVGSPPDRKPCBAGCFVCRQSKQ 360
Db 301 EKKRWGLIARYDFSLGKTFSLSHCYEGTVSFLPAQHTVGSPPDRKPCBAGCFVCRQSKQ 360

Qy 361 LEEOKKALYGLEAAEDVEEQVWVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 361 LEEOKKALYGLEAAEDVEEQVWVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420

Qy 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLEKGGKKRF 480
Db 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLEKGGKKRF 480

Qy 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGTEENPKPDHS 537
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGTEENPKPDHS 537

RESULT 4
US-10-618-941-121
; Sequence 121, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
```

Handwritten mark

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 121
; LENGTH: 537
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-618-941-121

Query Match      100.0%; Score 2888; DB 16; Length 537;
Best Local Similarity 100.0%; Pred. No. 3e-278;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLQSVLWVKQRCVSLPAPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60
Db 1 MGATGAAPLQSVLWVKQRCVSLPAPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60

Qy 61 EETDVHGKHQSGKQWKMEKPYAFTVHCVRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 120
Db 61 EETDVHGKHQSGKQWKMEKPYAFTVHCVRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 120

Qy 121 EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Db 121 EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180

Qy 181 YEINIDKYDGIWCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGIWCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGS 240

Qy 241 TDCVCYSTVGTSDAETSALHIVWGDLSLMDVSSVHHNSTLLRYSLVSLGFGYFGDIKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVWGDLSLMDVSSVHHNSTLLRYSLVSLGFGYFGDIKDS 300

Qy 301 EKKRWGLIARYDFSLGKTFSLSHCYEGTVSFLPAQHTVGSPPDRKPCBAGCFVCRQSKQ 360
Db 301 EKKRWGLIARYDFSLGKTFSLSHCYEGTVSFLPAQHTVGSPPDRKPCBAGCFVCRQSKQ 360

Qy 361 LEEOKKALYGLEAAEDVEEQVWVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 361 LEEOKKALYGLEAAEDVEEQVWVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420

Qy 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLEKGGKKRF 480
Db 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLEKGGKKRF 480

Qy 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGTEENPKPDHS 537
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGTEENPKPDHS 537

RESULT 5
US-09-969-896-11
; Sequence 11, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 562
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-969-896-11

Query Match      100.0%; Score 2888; DB 10; Length 562;
Best Local Similarity 100.0%; Pred. No. 3.2e-278;
```

Handwritten mark

Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLQSVLWVQKQRCVSLPARALLRRWRSPGAGAGADACSVPVSEIIAV 60
Db 26 MGATGAAPLQSVLWVQKQRCVSLPARALLRRWRSPGAGAGADACSVPVSEIIAV 85
Qy 61 EETDVHGKHGQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 120
Db 86 EETDVHGKHGQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 145
Qy 121 EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180
Db 146 EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 205
Qy 181 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGS 240
Db 206 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGS 265
Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 300
Db 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 325
Qy 301 EKKRWLGARYDFSGLTFTLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCROSKQ 360
Db 326 EKKRWLGARYDFSGLTFTLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCROSKQ 385
Qy 361 LEEBOKKALYGLEAAEDVEEWOVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 386 LEEBOKKALYGLEAAEDVEEWOVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445
Qy 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDDLKEGKKRF 480
Db 446 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDDLKEGKKRF 505
Qy 481 GHICSSHPSCCTVSNSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
Db 506 GHICSSHPSCCTVSNSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 562

RESULT 6
US-10-631-958-11
; Sequence 11, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631.958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-958-11

Query Match 100.0%; Score 2888; DB 16; Length 562;
Best Local Similarity 100.0%; Pred. No. 3.2e-278;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 86 EETDVHGKHGQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 145
Qy 121 EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180
Db 146 EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 205
Qy 181 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGS 240
Db 206 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGS 265
Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 300
Db 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 325
Qy 301 EKKRWLGARYDFSGLTFTLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCROSKQ 360
Db 326 EKKRWLGARYDFSGLTFTLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCROSKQ 385
Qy 361 LEEBOKKALYGLEAAEDVEEWOVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 386 LEEBOKKALYGLEAAEDVEEWOVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445
Qy 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDDLKEGKKRF 480
Db 446 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDDLKEGKKRF 505
Qy 481 GHICSSHPSCCTVSNSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
Db 506 GHICSSHPSCCTVSNSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 562

RESULT 7
US-10-315-597A-2
Sequence 2, Application US/10315597A
; Publication No. US20030162206A1
; GENERAL INFORMATION:
; APPLICANT: Sugiura, Masaaki
; APPLICANT: Kono, Keita
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It
; FILE REFERENCE: 02658CIP/HG
; CURRENT APPLICATION NUMBER: US/10/315,597A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: JP 2000-178039
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 537
; ORGANISM: Homo sapiens
US-10-315-597A-2

Query Match 99.7%; Score 2880; DB 14; Length 537;
Best Local Similarity 99.6%; Pred. No. 1.9e-277;
Matches 535; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGATGAAPLQSVLWVQKQRCVSLPARALLRRWRSPGAGAGADACSVPVSEIIAV 60
Db 1 MGATGAAPLQSVLWVQKQRCVSLPARALLRRWRSPGAGAGADACSVPVSEIIAV 60
Qy 61 EETDVHGKHGQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 120
Db 61 EETDVHGKHGQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 120
Qy 121 EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180
Db 121 EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180
Qy 181 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGS 240

```
QY 241 TDCVYSTVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFYGDIIDKS 300
DB 241 TDCVYSTVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFYGDIIDKS 300
QY 301 EKRWGLGARYDFSGLKTFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSKQQ 360
DB 301 EKRWGLGARYDFSGLKTFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSKQQ 360
QY 361 LEEEOKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
DB 361 LEEEOKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
QY 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKRF 480
DB 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKRF 480
QY 481 GHICSHSPCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
DB 481 GHICSHSPCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537

RESULT 8
US-09-784-810A-6
; Sequence 6, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-810A-6

Query Match 85.1%; Score 2456.5; DB 9; Length 471;
Best Local Similarity 97.5%; Pred. No. 2.4e-235;
Matches 459; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 78 MEKPYAFTVHCVRARRHRWKWAQVTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFI 137
DB 1 MEKPYAFTVHCVRARRHRWKWAQVTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFI 60
QY 138 NPFCKGKGRKRIYERKVAPLFTLASITTDII-----VTEHANOAKETLYEINID 186
DB 61 NPFCKGKGRKRIYERKVAPLFTLASITTDIIIGNKFYVNYVEVITEHANOAKETLYEINID 120
QY 187 KYDGI VCVGGDMGFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAGSTDVCY 246
DB 121 KYDGI VCVGGDMGFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAGSTDVCY 180
QY 247 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFYGDIIDKS 306
DB 181 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFYGDIIDKS 240
QY 307 GLARYDFSLKTFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSKQLEBEOK 366
DB 241 GLARYDFSLKTFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSKQLEBEOK 120
QY 187 KYDGI VCVGGDMGFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAGSTDVCY 246
DB 121 KYDGI VCVGGDMGFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAGSTDVCY 180
QY 247 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFYGDIIDKS 306
DB 181 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFYGDIIDKS 240
QY 307 GLARYDFSLKTFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSKQLEBEOK 366
DB 241 GLARYDFSLKTFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSKQLEBEOK 300
QY 367 KALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRKC 426
DB 301 KALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRKC 360

RESULT 10
US-09-784-810A-11
```

```
QY 427 SRFNFLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKRF 486
DB 361 SRFNFLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKRF 420
QY 487 HPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
DB 421 HPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 471

RESULT 9
US-10-876-281-6
; Sequence 6, Application US/10876281
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-876-281-6

Query Match 85.1%; Score 2456.5; DB 18; Length 471;
Best Local Similarity 97.5%; Pred. No. 2.4e-235;
Matches 459; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 78 MEKPYAFTVHCVRARRHRWKWAQVTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFI 137
DB 1 MEKPYAFTVHCVRARRHRWKWAQVTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFI 60
QY 138 NPFCKGKGRKRIYERKVAPLFTLASITTDII-----VTEHANOAKETLYEINID 186
DB 61 NPFCKGKGRKRIYERKVAPLFTLASITTDIIIGNKFYVNYVEVITEHANOAKETLYEINID 120
QY 187 KYDGI VCVGGDMGFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAGSTDVCY 246
DB 121 KYDGI VCVGGDMGFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAGSTDVCY 180
QY 247 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFYGDIIDKS 306
DB 181 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFYGDIIDKS 240
QY 307 GLARYDFSLKTFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSKQLEBEOK 366
DB 241 GLARYDFSLKTFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSKQLEBEOK 300
QY 367 KALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRKC 426
DB 301 KALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRKC 360
QY 427 SRFNFLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKRF 486
DB 361 SRFNFLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKRF 420
QY 487 HPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
DB 421 HPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 471
```

```
; Sequence 11, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; PRIOR FILING DATE: 2001-02-14
; PRIOR FILING DATE: 2000-02-14
; PRIOR FILING DATE: 2000-02-14
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-810A-11

Query Match      56.8%; Score 1640.5; DB 9; Length 326;
Best Local Similarity 96.3%; Pred. No. 2.9e-154;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 130 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178
Db 1 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVYVEVITEHANOAKE 60

Qy 179 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 238
Db 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120

Qy 239 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLLGYGFYGDIIK 298
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLLGYGFYGDIIK 180

Qy 299 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCQSK 358
Db 181 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCQSK 240

Qy 359 QOLEEEQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418
Db 241 QOLEEEQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300

Qy 419 DLILIRKSRNFLRLIRHTNQDQ 444
Db 301 DLILIRKSRNFLRLIRHTNQDQ 326

RESULT 11
US-09-969-896-2
; Sequence 2, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR FILING DATE: 2000-10-06
; PRIOR FILING DATE: 2000-10-06
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-896-2

Query Match      56.8%; Score 1640.5; DB 9; Length 326;
Best Local Similarity 96.3%; Pred. No. 2.9e-154;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 130 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178
Db 1 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVYVEVITEHANOAKE 60

Qy 179 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 238
Db 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120

Qy 239 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLLGYGFYGDIIK 298
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLLGYGFYGDIIK 180

Qy 299 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCQSK 358
Db 181 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCQSK 240

Qy 359 QOLEEEQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418
Db 241 QOLEEEQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300

Qy 419 DLILIRKSRNFLRLIRHTNQDQ 444
Db 301 DLILIRKSRNFLRLIRHTNQDQ 326

RESULT 12
US-10-631-958-2
; Sequence 2, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-958-2

Query Match      56.8%; Score 1640.5; DB 16; Length 326;
Best Local Similarity 96.3%; Pred. No. 2.9e-154;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 130 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178
Db 1 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVYVEVITEHANOAKE 60

Qy 179 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 238
Db 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120

Qy 239 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLLGYGFYGDIIK 298
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLLGYGFYGDIIK 180

Qy 299 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCQSK 358
Db 181 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCQSK 240
```


QY 359 QOLEBEQKALYGLAEAEVVEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418
DB 241 QOLEBEQKALYGLAEAEVVEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
QY 419 DLILIRKCSRNFNRLRFLIRHTNQDQ 444
DB 301 DLILIRKCSRNFNRLRFLIRHTNQDQ 326

RESULT 13
US-10-876-281-11
; Sequence 11, Application US/10876281
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-876-281-11

Query Match 56.8%; Score 1640.5; DB 18; Length 326;
Best Local Similarity 96.3%; Pred. No. 2.9e-154;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
QY 130 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178
DB 1 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANOAKE 60
QY 179 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPA 238
DB 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPA 120
QY 239 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLILGYGFYGDIIK 298
DB 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLILGYGFYGDIIK 180
QY 299 DSEKRWLGLARYDFSGLKTFLSHHCYEGTGVSLFPAQHTVGSFPRDKPCRAGCFVCROSK 358
DB 181 DSEKRWLGLARYDFSGLKTFLSHHCYEGTGVSLFPAQHTVGSFPRDKPCRAGCFVCROSK 240
QY 359 QOLEBEQKALYGLAEAEVVEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418
DB 241 QOLEBEQKALYGLAEAEVVEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
QY 419 DLILIRKCSRNFNRLRFLIRHTNQDQ 444
DB 301 DLILIRKCSRNFNRLRFLIRHTNQDQ 326

RESULT 14
US-09-784-810A-29
; Sequence 29, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08

; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 80432911
US-09-784-810A-29

Query Match 34.1%; Score 986; DB 9; Length 182;
Best Local Similarity 100.0%; Pred. No. 2e-89;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 AGAFGADACSVPVSEIIAIVETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHRWKWAQ 101
DB 1 AGAFGADACSVPVSEIIAIVETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHRWKWAQ 60
QY 102 VTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLA 161
DB 61 VTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLA 120
QY 162 SITTDIIIVTEHANOAKETLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNH 221
DB 121 SITTDIIIVTEHANOAKETLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNH 180
QY 222 PR 223
DB 181 PR 182

RESULT 15
US-10-876-281-29
; Sequence 29, Application US/10876281
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 80432911
US-10-876-281-29

Query Match 34.1%; Score 986; DB 18; Length 182;
Best Local Similarity 100.0%; Pred. No. 2e-89;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 AGAFGADACSVPVSEIIAIVETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHRWKWAQ 101
DB 1 AGAFGADACSVPVSEIIAIVETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHRWKWAQ 60
QY 102 VTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLA 161

Db	61	VTFWCPBEQLCHLWLQTLREMLEKLTSRPKHLLVFINPFGGKGQKRIYERKVAFLFLA	120
Qy	162	SITTDIIIVTEHANQAKETLYEINIDKYDGIVCVGGDGMFSEVLHGLIGRTORSAGVDQNH	221
Db	121	SITTDIIIVTEHANQAKETLYEINIDKYDGIVCVGGDGMFSEVLHGLIGRTORSAGVDQNH	180
Qy	222	PR 223	
Db	181	PR 182	

Search completed: September 3, 2005, 04:32:35
Job time : 115.269 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:25:44 ; Search time 99.8632 Seconds
(without alignment)
2753.633 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 NGATGAAPLOSVLVWVKQR.....QLVRLFARGIEENKPKDPSHS 537

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trenbl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2888	100.0	537	1 CEK1_HUMAN	Q8tct0 homo sapien
2	2428	84.1	531	1 CEK1_MOUSE	Q8k4q7 mus musculus
3	1904	65.9	409	2 Q62PK5	Q6zpk5 mus musculus
4	1811	62.7	339	2 Q6NX59	Q6nx59 homo sapien
5	1713	59.3	572	2 Q6GLV1	Q6glv1 xenopus lae
6	1495	51.8	485	2 Q6GMF3	Q6gmf3 xenopus lae
7	801.5	27.8	687	2 Q9VNA6	Q9vna6 drosophila
8	763	26.4	487	2 Q95S15	Q95s15 drosophila
9	666.5	23.1	410	2 Q7PRA8	Q7pra8 anopheles g
10	648	22.4	608	2 Q6USK2	Q6usk2 arabidopsis
11	619	21.4	700	2 Q6H6H1	Q6h6h1 oryza sativ
12	583.5	20.2	533	2 Q9LU45	Q9lu45 arabidopsis
13	576.5	20.0	532	2 Q6UZF6	Q6uzf6 homo sapien
14	567.5	19.7	549	2 Q9TZ11	Q9tzi1 caenorhabdi
15	480	16.6	586	2 Q949C3	Q949c3 oryza sativ
16	428	14.8	485	2 Q817L1	Q817l1 arabidopsis
17	417.5	14.5	1240	2 Q65419	Q65419 arabidopsis
18	369.5	12.8	763	2 Q91RB0	Q91rb0 arabidopsis
19	368.5	12.8	480	2 Q8H350	Q8h350 oryza sativ
20	349	12.1	732	2 Q9PHL3	Q9phl3 arabidopsis
21	343	11.9	280	2 Q6ZP59	Q6zp59 homo sapien
22	331.5	11.5	446	2 Q84S01	Q84s01 oryza sativ
23	330.5	11.4	757	2 Q94HY9	Q94hy9 oryza sativ
24	330.5	11.4	757	2 Q7KCS9	Q7kcs9 oryza sativ
25	323	11.2	685	2 Q7Q1P4	Q7q1p4 anopheles g
26	314.5	10.9	748	2 Q7XN57	Q7xn57 oryza sativ
27	310.5	10.8	641	2 Q9VY78	Q9vy78 drosophila
28	298.5	10.3	661	2 Q9VZW0	Q9vzw0 drosophila
29	291.5	10.1	458	2 Q14159	Q14159 schizosacch
30	271.5	9.4	654	1 SPH2_HUMAN	Q6ayb2 homo sapien
31	267.5	9.3	616	2 Q6AYB2	Q6ayb2 rattus norv

32	267.5	9.3	617	1 SPH2_MOUSE	Q9jia7 mus musculus
33	264.5	9.2	384	1 SPH1_HUMAN	Q9nyal homo sapien
34	264.5	9.2	384	2 Q96HV8	Q96hv8 homo sapien
35	264.5	9.2	398	2 Q96GK1	Q96gk1 homo sapien
36	264.5	9.2	470	2 Q8N632	Q8n632 homo sapien
37	260	9.0	388	2 Q88886	Q88886 mus musculus
38	257.5	8.9	624	2 Q86KF9	Q86kf9 dictyosteli
39	254	8.8	381	2 Q91ZN3	Q91zn3 mus musculus
40	254	8.8	382	2 Q8C115	Q8c115 mus musculus
41	254	8.8	504	2 Q88885	Q88885 mus musculus
42	253.5	8.8	579	2 Q74ZE3	Q74ze3 ashbya goss
43	248.5	8.6	687	2 Q06147	Q06147 saccharomyc
44	245.5	8.5	383	2 Q542F6	Q542f6 rattus norv
45	245.5	8.5	423	2 Q7JM91	Q7jm91 caenorhabdi

ALIGNMENTS

RESULT 1
CEK1_HUMAN
ID CEK1_HUMAN STANDARD; PRT; 537 AA.
AC Q8TCT0; Q9BYB3; Q9UGES;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ceramide kinase (EC 2.7.1.138) (Acylsphingosine kinase) (hCERK) (Lipid
DE kinase 4) (LK4).
GN Name=CERK; Synonyms=KIAA1646;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Leukemia;
RX MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;
RA Suglura M., Kono K., Liu H., Shimizugawa T., Minekura H., Spiegel S.,
RA Kohama T.;
RT "Ceramide kinase, a novel lipid kinase. Molecular cloning and
RT functional characterization."
RL J. Biol. Chem. 277:23294-23300(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Van Veldhoven P.P.;
RT "A search for lipid kinases."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA McClay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.C.T.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,

RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuayama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoso S., Murray J., Miller N., Mink P.,
RA Fulton R., Johnson D., Benis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.M., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelman L., Kim U.J., Shizuwa H., Simon M.I., Dumanek J.P.,
RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tiliahun Y., Wright H.;
RA "The DNA sequence of human chromosome 22.";
RA Nature 402:489-495(1999).
RN [4]
RP SEQUENCE OF 57-537 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21156230; PubMed=11258795;
RA Hirosewa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.;
RA "Identification of novel transcribed sequences on human chromosome 22
RT by expressed sequence tag mapping.";
RL DNA Res. 8:1-9(2001).
CC -1- FUNCTION: Catalyzes specifically the phosphorylation of ceramide
CC to form ceramide 1-phosphate. Acts efficiently on natural and
CC analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),
CC to a lesser extent on C2-ceramide and C6-dihydroceramide, but not
CC on other lipids, such as various sphingolipids.
CC -1- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.
CC -1- COFACTOR: Calcium and magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC -1- TISSUE SPECIFICITY: High level expression in heart, brain,
CC skeletal muscle, kidney and liver; moderate in peripheral blood
CC leukocytes and thymus; very low in spleen, small intestine,
CC placenta and lung.
CC -1- MISCELLANEOUS: Optimal pH is 6.0-7.5.
CC -1- SIMILARITY: Contains 1 DAKC domain.
CC -1- CAUTION: Ref.3 sequence differs from that shown due to erroneous
CC gene model prediction. An additional exon may exist between amino
CC acid positions 168 and 169.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB079066; BAC01154.1; -;
DR EMBL; AJ457828; CAD29884.1; -;
DR EMBL; AL096766; CAB62977.1; ALT SEQ.
DR EMBL; AL118516; -; NOT ANNOTATED_CDS.
DR EMBL; AB051433; BAB33316.1; -;
DR Genew; HGNC:19256; CERK.
DR GO; GO:000299; C:integral to membrane of membrane fraction; IDA.
DR GO; GO:0004685; F:calcium/calmodulin-dependent protein kinase. . . ; IDA.
DR GO; GO:0001729; F:ceramide kinase activity; IDA.
DR GO; GO:0000287; P:magnesium ion binding; IDA.
DR GO; GO:0006672; P:ceramide metabolism; TAS.
DR InterPro; IPR001206; DAKC.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011036; PH related.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAKGc; 1.
DR SMART; SM00046; DAKGc; 1.
DR SMART; SM00233; PH; 1.
KW Calcium; Kinase; Magnesium; Transferase.

FT DOMAIN 132 278 DAKC.
SQ SEQUENCE 537 AA; 59977 MW; 3DBFC0ED8D679P7F CRC64;
Query Match 100.0%; Score 2888; DB 1; Length 537;
Best Local Similarity 100.0%; Pred. No. 5e-233;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGATGAAPLQSLVWVKQRCVSLPAPARALLRWSPGFGAGAGADACSVPVSEIIAV 60
DB 1 MGATGAAPLQSLVWVKQRCVSLPAPARALLRWSPGFGAGAGADACSVPVSEIIAV 60
QY 61 EETDVHGKHGSGKWKQMEKPYAFTVHCIVRARHRHWKVAQVTFWCPBEOLCHLWLQTLR 120
DB 61 EETDVHGKHGSGKWKQMEKPYAFTVHCIVRARHRHWKVAQVTFWCPBEOLCHLWLQTLR 120
QY 121 EMLEKLTSRPKHLVFTNPFGKGQGRKRIYERKVAFLFTLASITTTDIIIVTEHANAOKETL 180
DB 121 EMLEKLTSRPKHLVFTNPFGKGQGRKRIYERKVAFLFTLASITTTDIIIVTEHANAOKETL 180
QY 181 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
DB 181 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
QY 241 TDCVCYSTVTGSDAETSAHIVVGDLSLAMDVSVVHNNSTLLRYSVSLLGTFYGDIIKDS 300
DB 241 TDCVCYSTVTGSDAETSAHIVVGDLSLAMDVSVVHNNSTLLRYSVSLLGTFYGDIIKDS 300
QY 301 EKKRWGLARYDFSGLKTFPLSHHCYEGTVSFLPAQHTVSGSPDRKPCRCACFCVCRQSKQ 360
DB 301 EKKRWGLARYDFSGLKTFPLSHHCYEGTVSFLPAQHTVSGSPDRKPCRCACFCVCRQSKQ 360
QY 361 LEBEQKALYGLEAAEDVEEWQVVGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
DB 361 LEBEQKALYGLEAAEDVEEWQVVGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
QY 421 ILIRKSRFNFLRLRHTNQOQDFTFVEVVRVKKQFTSKHMEDESDLEKGGKKRF 480
DB 421 ILIRKSRFNFLRLRHTNQOQDFTFVEVVRVKKQFTSKHMEDESDLEKGGKKRF 480
QY 481 GHICSSHPSCCCVTSNNSNCDGEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSSH 537
DB 481 GHICSSHPSCCCVTSNNSNCDGEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSSH 537
RESULT 2
CEK1_MOUSE STANDARD; PRT; 531 AA.
AC Q8K4Q7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ceramide kinase (EC 2.7.1.138) (Acylsphingosine kinase) (mCERK).
GN Name=Cerk;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;
RA Sugiura M., Kono K., Liu H., Shimizugawa T., Minekura H., Spiegel S.,
RA Kohama T.;
RA "Ceramide kinase, a novel lipid kinase. Molecular cloning and
RT functional characterization.";
RL J. Biol. Chem. 277:23294-23300(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart, and Thymus;
RX MEDLINE=22354683; PubMed=12456851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yananaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Nunata K., Okido T., Pavan W.J., Perte G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shinada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Vetrano R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 CC -!- FUNCTION: Catalyzes specifically the phosphorylation of ceramide
 to form ceramide 1-phosphate. Acts efficiently on natural and
 analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),
 to a lesser extent on C2-ceramide and C6-dihydroceramide, but not
 on other lipids, such as various sphingosines (By similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.
 CC -!- COFACTOR: Calcium and magnesium (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
 similarity).
 CC -!- TISSUE SPECIFICITY: High level expression in heart, brain and
 testis; low expression in spleen, liver and lung; not detected in
 skeletal muscle.
 CC -!- DEVELOPMENTAL STAGE: Highly expressed at 7 dpc and decreases
 rapidly thereafter.
 CC -!- SIMILARITY: Contains 1 DAGKc domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB079067; BAC01155.1; -;
 DR EMBL; AK042077; BAC31157.1; -;
 DR EMBL; AK052269; BAC34908.1; -;
 DR MGD; MGI-2386052; Cerk.
 DR GO; GO:0000299; C:integral to membrane of membrane fraction; ISS.
 DR GO; GO:0004685; F:calcium/calmodulin-dependent protein kinase. .; ISS.
 DR GO; GO:0001729; F:ceramide kinase activity; ISS.
 DR GO; GO:0000287; F:magnesium ion binding; ISS.
 DR GO; GO:0006672; P:ceramide metabolism; ISS.
 DR InterPro; IPR001206; DAGKc.
 DR InterPro; IPR011036; PH-related.
 DR Pfam; PF00781; DAGK_cat; 1.
 DR ProDom; PD005043; DAGKc; 1.
 KW Calcium; Kinase; Magnesium; Transferrase.
 FT DOMAIN 132 278
 FT CONFLICT 378 378 V -> M (in Ref. 2). BAC34908.
 FT CONFLICT 467 467 D -> Y (in Ref. 2). BAC34908.
 SQ SEQUENCE 531 AA; 59811 MW; 31PECS34C348AA0A CRC64;

Query Match

Best Local Similarity 84.1%; Score 2428; DB 1; Length 531;

Matches 450; Conservative 33; Mismatches 46; Indels 0; Gaps 0;

QY 1 MGATGAAPLOSVLVVKQORCAVSLPAPALLRWRSPGAGAPGACACSPVPSEITAV 60
 DB 1 MGAMGAAPLSVLVWVRRRCVSLPAPALLRWRSPGAGAPGADARSVLVSEITAV 60
 QY 61 BETDVHGQSGKQWQKMEKYAFVTVHCVKARHRWKWAQVTFWCPQEQLCHLWQLTLR 120
 DB 61 EEKDDCEGASSGRHWKMPFAFVHRVKVRHHRWKWARTVFWSADEQLCHLWQLTLR 120
 QY 121 EMLEKLSRPHLLVFINPFGKGQGGRIYERKVPAPLFTLASITTDIIIVTEHANOAKETL 180
 DB 121 GLLESLSRPHLLVFINPFGKGQGGRIYERKVPAPLFTLASITTEIIITEHANOAKETL 180
 QY 181 YEINIDKYGIVCGDGMFSEVLHGLIGRTORAGVDONHPRAVLVPSSLRIGIIPAGS 240
 DB 181 YEINTSDYGVIVCGDGMFSEVLHGLIGRTVQAGIDPNHPRAVLVPSTLRIGIIPAGS 240
 QY 241 TDCVCYSTVGTSDAETSAHIVGDSLAMDVSSVHHNSTLLRYSVSLGCGFYGDIIKDS 300
 DB 241 TDCVCYSTVGTNDATSAHIIIGDSLDAIDVSSVHYHTLLRYSVSLGCGFYGDIIKDS 300
 QY 301 EKKRWLGLARYDFSGKLTFLSHHCYEGTSLFPAQHTVGSPPDRKPCRAGCFVCRSQKQ 360
 DB 301 EKKRWMLGLRYDFSGKLTFLSHQYVEGTSLFPAQHTVGSPPDRKPCRAGCFVCRSQKQ 360
 QY 361 LEEQKALYCLEAAEDVEEHWVQVCGKFLAINATNMSCACRRSPGLSPAHLGDGSSDL 420
 DB 361 LEESEKALYCLEAAEVEEHWVQVCGKFLAINATNMSCACRRSPGLSPAHLGDGSSDL 420
 QY 421 ILIKRCRNFRLRLIRHTNQDQDFTFVBYVYKVFQFTSKEMEDSDSLKEGGKRF 480
 DB 421 ILIKRCRNFRLRLIRHTNQDQDFTFVBYVYKVFQFTSKEMEDSDSLKEGGKRF 480
 QY 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVHVHCOLVRLPARGIEE 529
 DB 481 GKICKDRPSCCTASRSSWNCDEVMHSPAIEVHVHCOLVRLPARGIEE 529
 RESULT 3
 Q6ZPK5 PRELIMINARY; PRT; 409 AA.
 ID Q6ZPK5
 AC Q6ZPK5
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE MKIAA1646 protein (Fragment).
 GN Name=MKIAA1646;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic tail;
 RX PubMed=14621295;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hirooka S.,
 RA Saka Y., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT III. the complete nucleotide sequences of 500 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:167-180(2003).
 DR EMBL; AK129416; BAC98226.1; -;
 DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
 DR ProDom; PD005043; DAGKc; 1.
 DR InterPro; IPR001206; DAGKc.
 DR Pfam; PF00781; DAGK_cat; 1.
 DR SMART; SM000046; DAGKc; 1.
 FT NON TER 1
 SQ SEQUENCE 409 AA; 45697 MW; 810E815B37E26808 CRC64;

Query Match 65.9%; Score 1904; DB 2; Length 409;

Best Local Similarity 87.6%; Pred. No. 9.2e-151;

```
Matches 352; Conservative 26; Mismatches 24; Indels 0; Gaps 0;
QY 128 SRPKHLVFNPGKGQGGKRIYERKVPAPLFTLASITTDIIIVTEHANQAKETIYEINIDK 187
Db 6 SRPKHLVFNPGKGQGGKRIYERKVPAPLFTLASITTEIIIVTEHANQAKETIYEINTDS 65
QY 188 YDGIYCVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPAGSTDVCVCS 247
Db 66 YDGIYCVCGDGMFSEVLHGLIGRTQQSAGIDPNHPRAVLPVSTLRIGIIPAGSTDVCVCS 125
QY 248 TVGTSDAETSAHIIHVGDGLAMDVSSVHHNSTLLRYSVLLGYGYGDIIDKSEKKRWLG 307
Db 126 TVGTNDAAETSAHIIIGDSLAIIDVSSVHHNSTLLRYSVLLGYGYGDIIDKSEKKRWLG 185
QY 308 LARYDFSLGTLFSLHHCYEGTVSFLPAQHTVGSPRDKPCRCAGFCVCRSQKQLEEEOKK 367
Db 186 LVRYDFSLGTLFSLHHCYEGTVSFLPAQHTVGSPRDNKPCRCAGFCVCRSQKQLEEEOKK 245
QY 368 ALYGLEAAEDVEEVQVVCGLFLAINATNMSCACRRSPRGLSPAHLGDSGSDLLIRKCS 427
Db 246 ALYGLEAAEDVEEVQVTCGKFLAINATNMSCACRRSPGGLSPAHLGDSGSDLLIRKCS 305
QY 428 RFNPLRLIRHTNQODQDFTFVEYVRVKFQFTSKHMEDESDLKSGGKKRFGHCSSH 487
Db 306 RFNPLRLIRHTNQODQDFTFVEYVRVKFPHFTSKHMEDESDLKSGGKKRFGHCICKDR 365
QY 488 PSCCCTVSNWNCDEVLHSPALEVRVHCOLVRLFARGIEE 529
Db 366 PSTCSASRWNCDGEVHSPALEVRVHCOLVRLFARGIEE 407

RESULT 4
Q6XN59
ID Q6XN59 PRELIMINARY; PRT; 339 AA.
AC Q6XN59;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE CERK protein.
GN Name=CERK;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX Strausberg R.;
RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC067255; AAH67255.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKc.
DR ProDom; PD05043; DAGKc; 1.
SQ SEQUENCE 339 AA; 37780 MW; A4C2ACDFF2B6F3D0 CRC64;
Query Match 62.7%; Score 1811; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.4e-143;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 MPSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPAGSTDVCVCSYVGTSDAETSA 258
Db 1 MPSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPAGSTDVCVCSYVGTSDAETSA 60
QY 259 LHTVWGDSLAMDVSSVHHNSTLLRYSVLLGYGYGDIIDKSEKKRWLGARVDFSLGKT 318
Db 61 LHTVWGDSLAMDVSSVHHNSTLLRYSVLLGYGYGDIIDKSEKKRWLGARVDFSLGKT 120
QY 319 FLSHHCYEGTVSFLPAQHTVGSPRDKPCRCAGFCVCRSQKQLEEEOKKALYGLEAAEDV 378
Db 121 FLSHHCYEGTVSFLPAQHTVGSPRDKPCRCAGFCVCRSQKQLEEEOKKALYGLEAAEDV 180
QY 379 EEWQVVCGLFLAINATNMSCACRRSPRGLSPAHLGDSGSDLLIRKCSRFNPLRLIRH 438
Db 181 EEWQVVCGLFLAINATNMSCACRRSPRGLSPAHLGDSGSDLLIRKCSRFNPLRLIRH 240
QY 439 TNQODQDFTFVEYVRVKFQFTSKHMEDESDLKSGGKKRFGHCSSHSPSCCCTVSNSS 498
Db 241 TNQODQDFTFVEYVRVKFQFTSKHMEDESDLKSGGKKRFGHCSSHSPSCCCTVSNSS 300
QY 499 WNCDEVLHSPALEVRVHCOLVRLFARGIEENPKPDHS 537
Db 301 WNCDEVLHSPALEVRVHCOLVRLFARGIEENPKPDHS 339

RESULT 5
Q6GLV1
ID Q6GLV1 PRELIMINARY; PRT; 572 AA.
AC Q6GLV1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC84197 protein.
GN Name=MGC84197;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
```

RESULT 6	
Q6GMF3	
ID	Q6GN
AC	Q6GN

Qy 164 TTDIIIVTEHANOAKETLYEINIDKYDGVCGDGMFSEVLHGLIGRTORSAGVDQNHPR 223
Db 191 CADVIVTEYANHARDNLYENLVKYGVCVGGDGFSEVLHGLIGRTORSAGVDQNHPR 250
Qy 224 AVLVPSSIRIGIIPAGSTDCVCYGTGTSDAETSAHLHIVGDSLAMDVSSVHHNSTLLRY 283
Db 251 AHLSCQINIRIGIIPAGSTDCIYATGINDPETSALHIIIGDQCPDLVSCVHHNRKRTFLKY 310
Qy 284 SVSLGGLGYGFIIDIKSEKKRWGLGARYDFSGSLKTLFSLHHCYEGTGFPLPAQHTVGSPRD 343
Db 311 SVSLGGLGYGFIIDIKSEKKRWGLGARYDFSGSLKTLFSLHHCYEGTGFPLPAQHTVGSPRD 370
Qy 344 RKPACRAGFCVCRQSQKLEEBEKKALYGLE-AAEDVEEVMQVCGKFLAINATNMSCACRR 402
Db 371 QTPCTSGCYICRQSSKQLEEQKQTVFGESEHRGQDDDDTTIKGRFMAINAVSMSCACPR 430
Qy 403 SPRGLSPAHLGDCSSDLILIRKSRNPLRFLIRHTNQDQDFTFVEYVRV 455
Db 431 TPKGLSPAHLADGSAADILVRKCSRDLDFLRHLIRHTSNKQDFPFVEYVRV 483
RESULT 7
Q9VNA6 PRELIMINARY; PRT; 687 AA.
AC Q9VNA6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG16708-PA (Cg16708-pb).
GN ORFNames=CG16708;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris L., Moshrefi A.,
RA Mount S.M., Moy M., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.M.,
RA Palazolo M., Pittman K.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington G., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RN Science 287:2185-2195(2000).
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirkas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence";
RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirkas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective";
RN Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RN Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RP SEQUENCE FROM N.A.
RX SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX EMBL: AE003603; AAF52040.2; -;
RG FlyBase; FBgn0037315; CG16708.
DR GO: 0048102; P: autophagic cell death; IEP.
DR GO: 0035071; P: salivary gland cell death; IEP.
DR InterPro: IPR001206; DAGKC.
DR Pfam: PF00781; DAGK cat; 1.
DR SMART: SM00046; DAGKC; 1.
SQ SEQUENCE 687 AA; 75690 MW; AAE081230A939412 CRC64;
Query Match 27.8%; Score 801.5; DB 2; Length 687;
Best Local Similarity 31.3%; Pred. No. 4.1e-58;
Matches 197; Conservative 97; Mismatches 206; Indels 129; Gaps 16;
Qy 11 QSVLWVQQRCAVSL--PARALLRWRSFGPAGAGPAGACASVPVSEIITAVETDVHGK 68
Db 59 QQLVWERLQIKOSPOGNEAKPL-----PPDSPAPQPGGICSYGQSHVLHLD--DVSI 111
Qy 69 HQSGGKQWKEP-----YAFTVHVCKRAR-----HRWKWAQ 101
Db 112 RSGDTKASSLKPPSPGSSRSCSGSDVAQKPTSQYLITINAMRLSKQSDTCNRWELRR 171
Qy 102 VTFWCPPEQLCHLWLQTLREMLEKLT---SRPKHLLVFINPFGKGQKRIYKKAAPLF 158
Db 172 LTFNSDPYIVRQWDELQIRLHSSSPTRMRVRLVFINPFGGRKAGAQTYERHVRPIF 231
Qy 159 TLASITTDIIIVTEHANOAKETLYEINIDKYDGVCGDGMFSEVLHGLIGRTORSAGVD 218
Db 232 QLAGVADATCITTORANQVNDILLSHDLGVYDAVCCVGGDGTVAEVINGLIFROMRELGLD 291


```
QY 219 QNHPRVLVPSSLRIGIIPAGSTDCVCTSVTSTDAETSAHIIHVGDLSLMDVSVSHNS 278
D 292 EQRPPIIRP-ALPVGVIIPAGSTDTIAYSMHGTADVRTAAIHVILGQHRGLDVCVSNQ 350
QY 279 TLLRYSVLGFGFGDIIDKSEKKRWGLGLARYDFSLKTLFSLHCHYGVTSFL----- 332
D 351 SLURFCASVLSYGLGDVNAQSENTRMGPRIYSGVNAFLNNVRLNNGYDAELRMLDEPDL 410
QY 333 ---PAQHTVGSPRD-----RKPCRAGCFVC-----RQSKQOLEEEOKKALYGL 373
D 411 LTTPLEDIPQSPDVSCLGESVPSVCYANCQRCSPASSIQEQRSLSLFTQESKEA---E 466
QY 374 AEDVE-----EWQVCGKFLAINAT 394
D 467 RNOQVETEDSHLAASEAALLRPRPRGNLRLPTGSISSMRNLGNQWKVRGNFFMCGA 526
QY 395 NMSACRSPRGLSPAHLGDGSSDLILIRKSRFNFLRLRHTNQ--DQDFTFVEVY 453
D 527 NITCARSNGISRYSHLGDCLDLILVKTSLNNVRLNNTAGRSGDIRNLPFVEVY 586
QY 454 RVKQFQTSKHMEDESDLKEGKKRFGHIC-----SSHPSCCTVSNSSWNCDE 504
D 587 RTRFRFRFTFSASEDYSLAGS-----CQITPPEMTAHS---STEFSSWNCDE 635
QY 505 VLHSPAIEVRVHCOLVRLPARGIENPKP 533
D 636 VVTDLDITMRSHCOLIEVFMGRPHSYSKP 664

RESULT 8
Q95S15
ID Q95S15 PRELIMINARY; PRT; 487 AA.
AC Q95S15;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE HL01539p.
GN ORFNames=CG16708;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacieb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY061001; RAL28549.1; -
DR FlyBase; FBgn0037315; CG16708.
DR GO; GO:0048102; P:autophagic cell death; IEP.
DR GO; GO:0035071; P:salivary gland cell death; IEP.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK cat. 1.
DR SMART; SM00046; DAGKc.1.
SQ SEQUENCE 487 AA; 53711 MW; A4E71EC40354BB07 CRC64;

Query Match 26.4%; Score 763; DB 2; Length 487;
Best Local Similarity 35.5%; Pred. No. 4.4e-55;
Matches 170; Conservative 73; Mismatches 146; Indels 90; Gaps 10;

QY 129 RPKHLLVFNPGGKGQKRIYERKVAFLFTASITTDIIIVTEHANOAKETLYEINIDKY 188
D 2 RVRLLVFNPGGKGAGATYERHVRPIFQLAGVDATCITQRANQVKDILLSHDLGVY 61
QY 189 DGIVCGGDMSEVLHGLIGRTQSAGVDQNHPRVLVPSSLRIGIIPAGSTDCVCTSVT 248
D 62 DAVCCVGGDGTVAEVINGLIIFQRMELGLDEQRPPIIRP-ALPVGVIIPAGSTDTIAYSM 120
```

```
QY 249 VGTSDAETSAHIIHVGDLSLMDVSVSHNSVHNTLLRYSVSLLOYGYFYGDIIKDKSEKKRWGL 308
D 121 HGTADVRTAAIHVILGQHRGLDVCVSNQGLRFLFCASVLSYGLDVAQSENRYRMGP 180
QY 309 ARYDFSLKTLFSLHCHYGVTSFL-----PAQHTVGSPRD-----RKPCRAGC 351
D 181 RRYISGVKAFLLNNGYDAELRMLDEPDLTTPLEDIQSPDVSCLGESVPSVCYANC 240
QY 352 FVC-----RQSKQOLEEEOKKALYGLAAEDVE----- 379
D 241 QRCSPASSIQEQRSLSLFTQESKEA---ERNQVETEDSHLAASEAALLRPRPRGNL 296
QY 380 -----EWQVCGKFLAINATNMSACRSPRGLSPAHLGDGSSDLILIR 424
D 297 LPTGSISSMRNLGNQWKVRGNFFMCGAIIITCARSNGISRYSHLGDCLDLILVK 356
QY 425 KCSRPNFLRLRHTNQ--DQDFTFVEVYVKFKFQTSKHMEDESDLKEGKKRFGHI 483
D 357 KTSLLNNVRLNNTAGRSGDIRNLPFVEVYTRFRFTFSASEDYSLAGS----- 408
QY 484 C-----SSHPSCCTVSNSSWNCDEGVLHSPAIEVRVHCOLVRLPARGIENPKP 533
D 409 CQITPPEMTAHS---STEFSSWNCDEGVVTDLDITMRSHCOLIEVFMGRPHSYSKP 464

RESULT 9
Q7PRA8
ID Q7PRA8 PRELIMINARY; PRT; 410 AA.
AC Q7PRA8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000010845 (Fragment).
GN Name=ENSANGG0000008356;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB0100859; EAA07556.2; -
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK cat. 1.
DR ProDom; PD005043; DAGKc.1.
DR NON_TER 1
FT NON_TER 410 410
SQ SEQUENCE 410 AA; 45669 MW; 79535E4C0D150154 CRC64;

Query Match 23.1%; Score 666.5; DB 2; Length 410;
Best Local Similarity 38.9%; Pred. No. 4.2e-47;
Matches 144; Conservative 59; Mismatches 140; Indels 27; Gaps 5;

QY 93 RHRWKAQVTFWCPBQOLCHLWLTREMLEKUTSRPKHLLVFNPGGKGQKRIYER 152
D 41 KTNWRVRAVALHNSEPRVELWYNRLSSDLRD-QNRPKHLLFLNPGGKNALALYER 99
QY 153 KVAFLFTASITTDIIIVTEHANOAKE--TLYEINIDKYDGVCGGDMSEVLHGLIGR 210
D 100 YAKFLFRLAGVDINLIITQRAQQIYDVTSKSLDNDYDGLVCCGGDGTFAELNGLVTR 159
QY 211 TORGAGVDQNHPRVLVPSSLRIGIIPAGSTDCVCTSVTSGTSDAETSAHIIHVGDLSLMD 270
D 160 TMDCGDIDKYP-AYLPKPNIPGIVPAGSTDTVACCLNGTDTIKTCIIHILGHSLD 218
```

```
Qy 271 VSVVH-----NSTLRYSVLLGYGYFDIIKDSEKRWGLARY 311
Db 219 ISAVYSADAAAKDEGASPGTGRPRPOLLLKFASALSYGYLGDIAVDSEKYRWGPKRY 278
Qy 312 DFSGLKTLFHHCEYGVTFPLPAQHTVGSPPDRKPCRAGCFVCRSQKQOLEEEOKKALYG 371
Db 279 DYSGFKKFLANRGYNAEIVVHLDRRGKQDPNDGVRCLDKCARCKAKYGRDCGGERASY- 337
Qy 372 LEAAEDVEEMVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFN 431
Db 338 ---EDDDTEPLVVRGKFLMWSGANISCSERSPOGSPFYCHLGDGLDLVLVRHTSMFN 394
Qy 432 LRFLIRHTNQ 441
Db 395 LRLLLTWTSK 404

RESULT 10
Q6USK2 PRELIMINARY; PRT; 608 AA.
AC Q6USK2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ceramide kinase.
GN Name=CERK;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22959668; PubMed=14563678; DOI=10.1101/gad.1140503;
RA Liang H., Yao N., Song J.T., Luo S., Lu H., Greenberg J.T.;
RT "Ceramide modulate programmed cell death in plants.";
RL Genes Dev. 17:2636-2641(2003).
DR EMBL; AY362552; AAQ62904.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKC; 1.
KW Kinase.
SQ SEQUENCE 608 AA; 68462 MW; 65EB9353692D9CA6 CRC64;

Query Match 22.4%; Score 648; DB 2; Length 608;
Best Local Similarity 31.6%; Pred. No. 2.5e-45;
Matches 182; Conservative 68; Mismatches 188; Indels 138; Gaps 19;

Qy 52 VPSVEIIAVERTDV-----HGKHGSGKQWKPKYAFTHVCVKRARRH--RWK 98
Db 71 IKFSDIYAVEFVSVGLVHSPKLGRLHAKCEFRLLNTQEMRYRTVHGFSQSPKEPCLWN 130
Qy 99 WAQVTFWCPBQLCHLWLQTLREMLEKLTSPKHLVFINFPFGKGQKRIYKRVAPLF 158
Db 131 LAATFTHGMDLQTCQWMDQLNYSLLIKEVERPNLLVFPKSGKSGKSWVE-TVSKIF 189
Qy 159 TLAITTDIIIVTEHANQAKETLYEI---NIDKYGVICVCGDGMFSEVLHG-LIGRTQ-- 212
Db 190 IRAKYNKVIIVTERAGHAFDVMASIQNKELTYDGIITAVGGDGFNEILNGLLSRLKVP 249
Qy 213 -----RSAG-----VDQ---NHPRAVLVPSL----- 231
Db 250 LPPSPDSFNSVQSGSSVPEPGDEVHTQKEHP---LLPDSVQEVMMFRTVNGSCE 306
Qy 232 -----RIGIIPAGSTDVCYSTVGTSDAFTSAHLHIVGDSLAMDVSSV--- 274
Db 307 GIEDPDHPFSERPRFGLIPAGSDAIWVCTTGARDPVTSAHLIILGRKPLDAMQVVRW 366
Qy 275 --HNNSTL---LRYSVLLGYGYFDIIKDSEKRWGLARYDPSGLKTLFSLHCHYEGTV 329
Db 367 KTAFTSTIEPIRYAASFAGYGYFDVISESEKRYWMMGPKRYDVGTGKIFLKHRSYAEV 426
```

```
Qy 330 SFLPAQ-----HT-----VGSPPDRK-PCRAGCFVCRSQKQOLEEEOKKALYGL 373
Db 427 MFEEAESENSKASLHTRSKTWPFRNTTTRSEKILCRANCKICNS-----KVGWNSASTTLN 481
Qy 374 AADVEEQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNLR 433
Db 482 PCPEKTRCWTCTGRFLTSIGAAMVMSNRNERAPDGLVVD AHLSDGFLHILIKDCSRPKYLW 541
Qy 434 FLIRHTNQ-ODQPDFTEVYVRVKFOFTSKHMEDESDLKEGKKRFGHICSSHPSCCC 492
Db 542 HLTELAKRGGEPLNFVEYHKTRAFITTS-----FG----- 573
Qy 493 TVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIE 528
Db 574 --EESVWNLDGEIFEAHQLSAQVLRLGLPLFASGPE 607

RESULT 11
Q6H6H1 PRELIMINARY; PRT; 700 AA.
AC Q6H6H1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative ceramide kinase.
GN Name=P0519E06.23; Synonyms=QJ1003 B06.39;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
clone:P0519E06.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005006; BAD25678.1; -.
DR EMBL; AP004676; BAD25337.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR ProDom; PD005043; DAGKC; 1.
KW Kinase.
SQ SEQUENCE 700 AA; 77961 MW; FBB2C2A3DF31C75D CRC64;

Query Match 21.4%; Score 619; DB 2; Length 700;
Best Local Similarity 28.3%; Pred. No. 8.2e-43;
Matches 191; Conservative 86; Mismatches 192; Indels 206; Gaps 25;

Qy 1 MGATGAAPLQSVLWVKKQCAVSLEP-ARALLRWRRSPGAGAGADACSPVPSIIA 59
Db 70 LGQHGDIEEVSSSCW-----SSIIMQPKLESKLKF-----SDVVAVELLEVP 112
Qy 60 VEE-----TDVHGKHGSGKQWKPKYAFTHVCVKRARRHWK--AQVTFWCPBQLC 112
Db 113 VCEPWNARATVQGGKIN-----TEMNR---FVIHTVTRPKRPSWPVCEYIFGHKDOQTC 164
Qy 113 HLWLQTLREMLEKLTSPKHLVFINFPFGKGQKRIYKRVAPLFTLASITTDIIIVTEH 172
Db 165 KTWVEHKTCTNCKEQRPKSLMVFVHPLCGKGRCKNWE-TVAPLFEAKVTKVIQTOR 223
Qy 173 ANQAKETLYEI---NIDKYGVICV----- 194
Db 224 AGHAYDTLASLSDKDLKKFDGVIANTINACLSLFDIKHNYKMSARPENTLSYDPSAA 283
Qy 195 -----GGDGMFSEVL-----HGLIGRTORSA 215
```

```
Db 284 SGHKSMLFYCFIINMKQEHNRNDLSNSELTDGDANAISSSNTPTDDHBPILLSTRST 343
Qy 216 GVD-----QHPRVAVL--PSS-LRIGIIPAGSTDC-----VCYSTVGTSDAE 255
Db 344 GLD1SSSDSPFCNGDQVPLVFFNNWFLGIIIPSGSTDAIVLSPDVC-STTGERDPV 402
Qy 256 TSALHIVGDSLAMDVSSV-----HNSLLRAYSVSLGFGYFGVDIIKDSKKRWLG 307
Db 403 TSALLIILGRRISLDIAQVVRKWSPSAEVSTVRYAASFAGYFGEVIRESEKYRWG 462
Qy 308 LARYDFSGLKTLFSLHHCYEGTVSFLPAQHT-----VGSPPDRKP-CR 348
Db 463 PARYDFSGTWFLKHSYEAKVAFLNGWTHSLTASAENNANGVQTLQVHQNRHKTICR 522
Qy 349 ACCFVCR---QSQQLBEOKALYGLEAAEDVEBQVVCVGFPLAINATNMSCACRRSPR 405
Db 523 TNLCLCKGTSTSEQSEDENPDS---SRTACETPKVWMSKGRFLSVGAIVSCRNERAPD 579
Qy 406 GLSPAHLGDSGLLILIRKCSRFNLFRLIRHTNQ-----QDQDFTFVEYVRKPKQF 460
Db 580 GLUVAHLSDGFLHULLIRDCP-----LPYLMHLTQFTYKKGSDPLSFVHHKTOAFTF 635
Qy 461 TSKHMEDESDILKEGKKRFGHCISHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLV 520
Db 636 ISSHDE-----SVNLDGELLQACEVSVQAFRGIV 665
Qy 521 RLFGARGIE--ENPKP 533
Db 666 NLFASGPEKMENTAP 680

RESULT 12
Q9LU45 PRELIMINARY; PRT; 533 AA.
AC Q9LU45;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similarity to unknown protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:31-63(2000).
DR EMBL; AB023044; BAA97392.1; -.
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR ProDom; PD005043; DAGKC; 1.
SQ SEQUENCE 533 AA; 60153 MW; 1B851C7606B03E0E CRC64;

Query Match 20.2%; Score 583.5; DB 2; Length 533;
Best Local Similarity 33.5%; Pred. No. 5.3e-40;
Matches 158; Conservative 57; Mismatches 157; Indels 99; Gaps 16;

Qy 52 VPVSIIVAEETDV-----HGKHQSGKQKMEKPYAFTVCHVKGRARRH--RWK 98
Db 71 IKFSIYAEVFSYGLVHSPKLGRLHAKCEFRERLLNTQEMVFTVHGFQSPKEPCLWN 130
Qy 99 VAQVTFWCPPEQLHLWLQTLREMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLF 158
Db 131 LAAFTFGHMDLQTCQSWMDQLNYSLIKVERPRNLVLFVHPKSGKNGSKVWE-TVSKIF 189
Qy 159 TLASITTDIIVTEHANQAKETIYEI---NIDKVDGIVCVGGDMFSEVLHG-LIGRTO-- 212
```

```
Db 190 IRAKNTKVIIVTERAGHAFDVWASIQNKELTYDGIITAVGGDGFNEILNGYLLSRKVP 249
Qy 213 -----RSAG-----VDQ--NHPRVILVPSL----- 231
Db 250 LPPSPDSFNSVQSGRSSVPPEGDVHETDQKEHP---LLPDSVQVWMNFRIEDPDHP 306
Qy 232 -----RIGIIPAGSTDCVYSTVGTSDAETSALHIVVGDSLAMDVSSV-----HNSL 280
Db 307 FSSRPFRPLIPAGSTDAIVMCTTGARDPVTLSALHILGRKFLDAMQVVRWKTASTSTI 366
Qy 281 ---LRYSVSLGFGYFGDIIKDSKKRWGLARYDFSGLKTFSLHHCYEGTVSFLPAQ-- 335
Db 367 EPIRYAASFAGYFYGVDISESEKYRWGPKRDIYGTGKIFLKRSEYAEVMEFEASE 426
Qy 336 ---HT-----VGSPPDRK-PCRAGCFVCRCQKQLEEBEQKALYGLEAAEDVEW 381
Db 427 NSKASLHTRSKTWPFRNTTRSEKILCRANCKICNS-----KVGWNSATTTLNPCPEKTRW 481
Qy 382 QVVCGLKFLAINATNMSCACRRSPRLGSPAHLGDSGLLILIRKCSRNFPL 432
Db 482 CRTKGRFLSIGAAVMSNRNERAPDGLVVDHLSDFLHLILIKDCSRPKYL 532

RESULT 13
Q6UZF6 PRELIMINARY; PRT; 532 AA.
AC Q6UZF6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ceramide kinase-like protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14681825;
RA Tuson M., Marfany G., Gonzalez-Duarte R.;
RA "Mutation of CERKL, a novel human ceramide kinase gene, causes
RT autosomal recessive retinitis pigmentosa (RP26).";
RL Am. J. Hum. Genet. 74:128-138(2004).
DR EMBL; AY357073; AAR13670.1; -.
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKC; 1.
KW Kinase.
SQ SEQUENCE 532 AA; 59602 MW; C73E590F7C25EED1 CRC64;

Query Match 20.0%; Score 576.5; DB 2; Length 532;
Best Local Similarity 27.6%; Pred. No. 2.1e-39;
Matches 149; Conservative 102; Mismatches 195; Indels 93; Gaps 12;

Qy 29 RALLRWRSPPGAGAPGACACSVPSSE-----IIAVEE----- 62
Db 45 RILLR-----GIFEIGRSDCVLSEALRWPIQPERPAGDSKYDLLCKEERFLKD 97
Qy 63 -TDVHGKHQSGKQKMEKPYAFTVCHVKGRARRHVKWQVTFWCPEQLCHLWLQTLRE 121
Db 98 IFSVKLRCSRCSVKQQRSGTLGITLFLCLCKEONKLNSTDLINLSHDCDIWFRQPKK 157
Qy 122 MLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIVTEHANQAKETLY 181
Db 158 ILAGFNRPKSLKTLNLPQSHKKEATQYVEKVEPLKLAGIKTDVTIMBEGHALSLK 217
Qy 182 EINTDKYDGVICVGGDMFSEVLHGLIGRTQSRAGVQNHPRVAVLP--SSLRIGIIPAG 239
Db 218 ECELQGFVGVCVGGDSASEVAHALLLRQAKNAGMETDR---ILTPVRAQLPLGLIPAG 274
```

[illegible]

RESULT 14

ID	Q9TZ11	PRELIMINARY:	PRT;	549 AA.
AC	Q9TZ11;			
DT	01-MAY-2000	(TrEMBLrel. 13, Created)		
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)		
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)		
DE	Hypothetical protein T10B11.2.			
GN	Names=T10B11.2; ORFNames=T10B11.2;			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Peloderinae; Caenorhabditis.			
OX	NCBI TaxID=6239;			

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RT WormBase Consortium;
RT "Genome sequence of the nematode *C. elegans*: a platform for
RL investigating biology. The *C. elegans* Sequencing Consortium.";
RL Science 282:2013-2018(1998).

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Minx P., Kemp K.;
RT "The sequence of *C. elegans* cosmid T10B11.";
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.

AN
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.

[3]

AN
 RP
 RC
 RA
 RC
 RL
 [a]
 SEQUENCE FROM N.A.
 STRAIN=Bristol N2;
 Wilson R.;
 Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.

DR *EMBL*, *AF262353*; *AKC87460.1*,
DR *PIR*, *T33517*, *T33517*.
DR *WormBase*; *WBGene0020398*; *T10B11.2*.
DR *WormPep*; *T10B11.2*; *CE18241*.
DR *InterPro*; *IPR001206*; *DAGKC*.
DR

DE95737555534EEB CRC64;
SQ SEQUENCE 549 AA; 62425 MW;
KW Hypothetical protein.

Query Match 19.7%; Score 567.5; DB 2; Length 549;

	Best Local Similarity	30.8%; Pred. No. 1.2e-38;	
	Matches	139; Conservative	91; Mismatches 160; Indels 61; Gaps 13;
Qy	89	VKRARRHRWKWAO--VTFMCPEEQCHLWLQTILREMLEKLTSPXGHLLVFINPFGKGQG	146
Dd	123	VYKKDQKQKLKQIPVIFFYTSSR--DYWHSLDDTTLRRVKNPKNIIIFINPFGNGKA	180
Qy	147	KRIYERKVAPLFLLA-SITTDIIIVTHANQAETLYEINIDKY---DGIVCVGGDGMFSE	202
Dd	181	QKIFKDNVDFAFWLTPGLRYKAVVLTERANHARDYIIVEMPPEDQWSAIDGLYSVGDDGLFNE	240
Qy	203	VLHCLIGRTORSAGVDONHPRAVLVPSSLIRIGIIPAGSTDVCVCTSVGTSTDAETSALHIV	262
Dd	241	LLSGALLRTQTDAGRINDPSLSHLVTPHIRFGIIGAGSANSIVSTVHETNDHTATSVAHIA	300
Qy	263	VGDSLAMDVSVVHNNTLLRYSVSLICGYFYGDIIKDSEKKRWLGRLARYDFSGLTPLSH	322
Dd	301	IGSECNVDVCTVHQHKLIIRISANAISYGWLGDVRDSEBYRCGLGPRIYQMSALRTIIRH	360
Qy	323	HCEYGVTSFLPAOHTVGSPRDR-KPCRAGCFVCQRQSKQQLLEBQKALKYLEAAEDVEEW	381
Dd	361	PIYRGMVQFSLSHKENVNPKDLPPCLEPCVCMK-----POGNDRKYDHHWA-----EFT	411
Qy	382	QVCGKFLAINATNMSCACRRSPRGSLSPAHLHGSGSDLLILIKCSRFNFRLFRLRHRT--	439
Dd	412	HVICCVIPTVTFP-----TPYGLAFPTGIIGDTLDLALVPRISFRHNNQPMRKVMY	463
Qy	440	--NQDQDFTFVEVYRVKKFQFTSKHMEDEDSDLKEGGKKRFGHICSSHPSCCCTVSNS	497
Dd	464	GKQLYELD-PSLNCYRVTKWSY-----QPDAQEDPG-----	495
Qy	498	SWNCDEVLHSP---AIEVRVHCQLVRLPAR	525
Dd	496	VWMLDGEILEQPKDEPLHFKLHPQLSIFFGR	526

RESULT 15

QD	Q949C3	PRELIMINARY;	PRT;	586 AA.
AC	Q949C3;			
DT	01-DEC-2001	(TREMBLrel. 19, Created)		
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2004	(TREMBLrel. 26, Last annotation update)		
DE	Hypothetical protein C875ERIPDS.			
GN	C875ERIPDS;			
OS	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzoaeae; Oryza.			
NC	NCBI_TaxID=4530;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21329048;	PubMed=11435398;	DOI=10.1101/gr.GR-16117R;	
RA	Mayer K., Murphy G., Tarchini R., Wambutt R., Volkkaert G., Pohl T.,			
RA	Diesterhoef A., Stiekema W., Entian K.D., Terry N., Lemcke K.,			
RA	Haase D., Hall C.R., van Dodeweerd A.M., Tingey S.V., Mewes H.W.,			
RA	Bevan M., Bancroft I.;			
RT	"Conservation of microstructure between a sequenced region of the			
RT	genome of rice and multiple segments of the genome of Arabidopsis			
RT	thaliana.";			
RL	Genome Res. 11:1167-1174(2001).			
DR	EMBL; AJ307662; CAC39069.1; -.			
DR	Gramene; Q949C3; -.			
DR	GO; GO:0004143; F-diacylglycerol kinase activity; IEA.			
DR	GO; GO:0007203; P:protein kinase C activation; IEA.			
DR	InterPro; IPR001206; DAGKC.			
DR	ProDom; PD005043; DAGKC; 1.			
KW	Hypothetical protein.			
SO	SEQUENCE	586 AA;	64892 MW;	1C45DFB3670E1E22 CRC64;

Query Match	16.6%;	Score 480;	DB 2;	Length 586;
Best Local Similarity	26.8%;	Pred. No. 2.8e-31;		
Matches 152; Conservative	81;	Mismatches 173;	Indels 162;	Gaps 21;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 6, 2005, 07:01:07 ; Search time 5163.87 Seconds
(without alignments)
3958.373 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 MGATGAEPLOSVLWVKQOR.....QLVRLFARGIENPKPDHS 537

Scoring table:

```
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
```

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10631958/runat_02092005_165816_4562/app_query.fasta_1.1941
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODES=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CGN 1 1 9235 @runat_02092005_165816_4562 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2433.5	84.3	1824	3 AK042077	Mus muscu
2	2424.5	84.0	4248	3 AK052269	Mus muscu
3	1518	52.6	1063	4 BM479389	AGENCOURT
4	1367	47.3	1059	5 BQ054406	AGENCOURT
5	1321	45.7	1078	5 BQ063738	AGENCOURT
6	1299.5	45.0	797	7 CK000755	AGENCOURT
7	1284	44.5	1047	5 BQ057191	AGENCOURT
8	1275	44.1	713	7 CK000525	AGENCOURT
9	1260	43.6	820	6 CD855311	AGENCOURT

10	1244	43.1	732	7	CF135528	UI-HF-BN0
11	1205	41.7	758	6	CB246749	UI-M-F10-
12	1181	40.9	653	5	BX952302	DKF2p781L
13	1170	40.5	661	7	CN296312	170005321
14	1053.5	36.5	812	7	CK603033	AGENCOURT
15	1012	35.0	581	5	BP224560	BP224560
16	1003	34.7	1017	5	BO879739	AGENCOURT
17	997	34.5	584	5	BP310011	BP310011
18	986	34.1	779	5	BU265377	603374309
19	982	34.0	584	5	BP309990	BP309990
20	971	33.6	763	7	CK364327	AGENCOURT
21	968	33.5	573	7	CF138275	UI-HF-BN0
22	964	33.4	819	7	CN235868	WLB086C10
23	959	33.2	550	7	CF138634	UI-HF-BN0
24	943.5	32.7	759	5	BU264162	603814452
25	942	32.6	541	7	CN296311	170005321
26	925	32.0	725	5	BU214294	603756937
27	924	32.0	607	6	CB581157	AMGNNUC.N
28	916	31.7	576	6	CB608292	AMGNNUC.N
29	916	31.7	582	5	BP274786	BP274786
30	914	31.6	653	6	BY754042	BY754042
31	909	31.5	608	6	CB580936	AMGNNUC.N
32	876	30.3	553	6	CA578972	K0731A08-
33	873	30.2	674	1	AJ739034	AJ739034
34	871	30.2	550	2	AW499858	UI-HF-BN0
35	849	29.4	902	5	BU373554	603811294
36	838	29.0	545	6	CB611947	AMGNNUC.N
37	829	28.7	662	2	BB638867	BB638867
38	828	28.7	488	5	BU430459	UI-HF-BN0
39	813	28.2	540	6	CB612897	AMGNNUC.N
40	811	28.1	474	2	AW503999	UI-HF-BN0
41	806	27.9	638	1	AJ739022	AJ739022
42	805	27.9	449	2	AW503483	UI-HF-BN0
43	797.5	27.6	759	5	BU290807	604162346
44	796	27.6	656	6	BY739924	BY739924
45	786.5	27.2	680	2	BB235873	BB235873

ALIGNMENTS

RESULT 1	AK042077	1824 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK042077	Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630056D11 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence.			
ACCESSION	AK042077	GI:26334912			
VERSION	AK042077.1	HTC; CAP trapper.			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,				

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Iihikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multipillar sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

TITLE

JOURNAL Functional annotation of a full-length mouse cDNA collection
 MEDLINE Nature 409, 685-690 (2001)

REFERENCE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

TITLE

JOURNAL Analysis of the mouse transcriptome based on functional annotation
 MEDLINE Nature 420, 563-573 (2002)

REFERENCE

6 (bases 1 to 1824)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hangaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M., and Hayashizaki, Y.
 Direct Submission

TITLE

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216]

COMMENT

CDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
 Trust/MRC building Addenbrookes Hospital Cambridge) whose
 assistance we gratefully acknowledge.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/.

FEATURES

source
 1. 1824
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:A630056D11"
 /db_xref="taxon:10090"
 /clone="A630056D11"
 /tissue_type="thymus"
 /clone_libs="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="3 days neonate"
 85. -1680

CDS

note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN
 SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED)
 PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTFQ9UGES,
 evidence: FASTY, 86.5%ID, 100%length, match=945)
 putative"
 /codon_start=1
 /protein_id="BAC31157.1"
 /db_xref="GI:26334913"
 /translation="MGAMGAAEPLHSLVWVKRRCAVSLSEPARALLRWRRSPPEGPSA

PGADARSLVSEIIIAVEEKDDCEKHAASSGRWHKVENPFAFTVHRVKRVRHHRWKARV
 TFSAEQDLCHLWLTGLLESITSRPKHLVFINPFGKGQCKRIYKTYKVAFLFTL
 ASITTEIIITEHANOAKETUIYNTSDGIVCVGGDMFSEVLGHVIGRTQOSAGID
 PNHPRALVPSTLRIGIIPAGSDTCVCYSTVGTNDAAETSAHIIIGDSLADYSSVHV
 HNTLLRYSVSLKGTGPGFDLTKDSEKRWMLVRFDSGLKTLFSLHGYEGTSLSFLEA
 QHTVGSFRDNKPCRAGCFVCRSQKQLEEEKALYGLNAAEEMEEQVTCGKFLAIN
 ATNMSCAPRSPGSLPFAHFGSSDLILIRKCSRNFRLFLIRHNTQDQDFDTFV
 EVYRVKHFHTSKHVEDNDNSKEQKQFKGKICKDRPSCTSASRSSMNCDEGVMS
 PAIEVRVHCQLVRLFARGIEES"

ORIGIN

Alignment Scores: Length: 1824
 Pred. No.: 8.32e-231 Matches: 453
 Score: 2433.50
 Percent Similarity: 90.19% Conservative: 34
 Best Local Similarity: 83.89% Mismatches: 48
 Query Match: 84.26% Indels: 5
 DB: 3 Gaps: 1
 US-10-631-958-10 (1-537) x AK042077 (1-1824)
 Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
 Db 85 ATGGGGCAATGGGGCGCGGAGCCCTGCCTCCTGCTGGTGGTGAACCGCGACGC 144
 Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
 Db 145 TGTGCGCTCAGCTGGAGCGCGCGGGCTCTGCTACCTGGTGGCGGAGCCCGAGCC 204
 Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluLeuLeuAlaVal 60
 Db 205 GGGCCCTCCGCGCGGTCGCGATGCCGCTCGGTACTGCTGTCGAGATCATCGCTGTT 264
 Qy 61 GluLthrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
 Db 265 GAGGAAAACAGCAGCTCGGAGAACACGCTCCAGTGGCGATGGCTATAAATGGAAAT 324
 Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100
 Db 325 CCGTTCGATTCAGTCCACCGTGTGAGCGAGTACGACATCACCGCTGGAAGTGGCA 384
 Qy 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
 Db 385 CGGTGACCTCTTGGAGCGCGCGAGCAGCTGTGTCCACCTGTGGCTGCAGACCTCGT 444
 Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheLeuAspProPhe 140
 Db 445 GGGTGTCTGGAGACCTTGACTTCAAGACCGAGCAGCTTGTCTGGTATTTCATCAACCTTC 504
 Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
 Db 505 GGAGGAAAGGTGAGGCAAGCGCATCTATGMAAAAACAGTGGCGCTCTGTTTACCTTG 564
 Qy 161 AlaSerIleThrAspIleLeuValThrGluHisAlaLeuGlnAlaLysGluThrLeu 180
 Db 565 GCTTCCATCATACCGAGATCATCATTTACAGAGCATGCCAACCAAGCAAGGAGACATTTA 624
 Qy 181 TyrGluLeuAsnIleAspLysTyrAspGlyLeuValCysValGlyGlyAspGlyMetPhe 200
 Db 625 TACGAGATCAACACAGACAGCATATGATGGCATCTGTGCTAGGTGGGAGCGGATGTTT 684
 Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
 Db 685 AGCGAGGTGTCATGGGTGATTGGGAGGAGCGCAGCAGAGCGCTGGTATCGACCCCAAT 744
 Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleLeuProAlaGlySer 240
 Db 745 CACCCCGAGCGGTGCTGGTGCCTCCAGTACCTCCAGGATCGGCATCATACCGCAGGTGCC 804
 Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
 Db 805 ACAGATTGTGTGTGTACTCAACAGTGGGCACAAACGACGAGAGACATCGGCTTTGCAC 864

/clone.lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="13 days embryo"

CDS

260_11855
/note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN
SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED)
PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTP [Q9UGES,
evidence: FASTY, 86.5%ID, 100%length, match=945)
putative"
/codon_start=1
/protein_id="BAC34908.1"
/db_xref="GI:26342492"
/translation="MGAMGAAEPLHSLVWVRRRCVLSLEPARALLRWRSPBPGPSA
PGADARSLVSEIIAVEKDDCEKHSAGRWKHMENFAFTVHRVKRVRHRRWKWARV
TFWASDQLCHLWLOTURGLLESLSRPKLLVFNPFPGKGQGRKYEXTVAPLFTL
ASITTEIITEHANAOKETLYELNDSYDGI VCVGGDMPSEVLHGVIGRTQOSAGID
PNHPRVLVSTLRIGLIPAGSDVCVSTVGNDAETSALHIIIGDSLDAIDVSSVHY
HNTLLRYSLLGFGVDLIKOSKKRWMLVRYDFSGUKTFLSHOYEGTSLFPLA
QHTVGSPRDNCRAGFCVQSRQKLEEBEKALYGLENAEBEEMQVTCGKFLAIN
ATNNSCAPRS PGLSPFAHLGDGSDLIILIRKCSRFNFLRIHTNQDQDFTFV
EVRVKPFHTSKHVEVDNDSKEQEKFKGKICKDRPSC TCSASRSSMNCDEVMHS
PAIEVRVHCQLVRLFARGIEEBS"

ORIGIN

Alignment Scores:

Pred. No.: 2,42e-229 Length: 4248
Score: 2424.50 Matches: 452
Percent Similarity: 90.00% Conservative: 34
Best Local Similarity: 83.70% Mismatches: 49
Query Match: 83.95% Indels: 5
DB: 3 Gaps: 1

US-10-631-958-10 (1-537) x AK052269 (1-4248)

```
Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
Db 260 ATGGGGGCNAATGGGGGGCGGAGCCGCTGCACCTCCGCTGCTGGGTGAAACGGCGACGC 319
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 320 TGTGCGGTACAGCTGGAGCCGCGGGGCTCTGTACGCTGTGTGGCGAGCCCGGAGCCC 379
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db 380 GGGCCCTCCGGCGGGTCCGATGCCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 439
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 440 GAGGAAAAGACGATCGCGAGAAACACGCTCCAGTGGCCGATGGCATATAAATGGAAAT 499
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100
Db 500 CCGTTCGATTCACAGTCCACCGTGTGAGCGAGTACGACATCACCGCTGGAAGTGGGCA 559
Qy 101 GlnValThrPheTrpCysProGluGlnGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db 560 CGGGTGACCTTCTGGAGCGCCGAGCAGCAGTGTGTCACTGTGGCTGCAGACCTCCGT 619
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 620 GGGTGTCTGGAGACCTTGACTTCAAGACCGAGCAGTGTGTGGTATTATCAACCCCTTTC 679
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 680 GGAGGAAAGGTCAAGCAACGCGATCTATGAAAAAACAGTGGCGCTCTGTATTACCTTG 739
Qy 161 AlaSerIleThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 740 GCTTCCATCACTACGGAGATCATATTACAGAGCATGCCAACCAAGCAAGGAGACTTGA 799
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 800 TACGAGATCAACACAGACAGCATATGATGGCATGTGTGCTGCTGCTGCTGCTGCTGCTGCT 859
```

```
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 860 AGCGAGGTGCTGATGGGGTGAATTGGAGGACGAGCAGAGCGCTGTGTATCGACCCCAAT 919
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 920 CACCCCGGACCGTGTGGTGGCCAGTACCTCAGGATCGCATATACCGCAGGCGTCC 979
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 980 ACAGATTGTGTGTACTCAACAGTGGGCAACACGACGACAGACATCGGCTTTGCAC 1039
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
Db 1040 ATCATATTGGGGACTCAGTGGCAATAGACGTCTCTCTGTGCATACCATACACGCTG 1099
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
Db 1100 CTGCGGTACTCGGTTTCTCTGTGGGTACGGTTCTACGGGACTTAATCAAGGACAGT 1159
Qy 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
Db 1160 GAAAAGAACGGTGGATGGGCTCTCGGTATGATTTCTCAGGGTTGAAGACCTTTCTC 1219
Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db 1220 TCTCATCAGTACTATGAAGGACACATGTCTCTCTCCAGCAGACACGCTGGGATCT 1279
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db 1280 CCACGGGACAAATAAACCTCCCGGCTGGGTGTCTGTGTGCAGGACAGCAACACAG 1339
Qy 361 LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db 1340 CTGAAGAGAGAGAGAGAAAGCCCTGTATGGCTGGAGAACGCCGAGGAAATGGAAGAG 1399
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db 1400 TGGCAAGTGACATGTGGGAAGTTCTTGGCCATCAATGCCACCAACATGTCTGTGCTGT 1459
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db 1460 CCTCGGAGCCCTCGGGGCTGTCCCATTTGCCCATCTGGGAGATGGGTCTCTCGACCTC 1519
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 1520 ATCTTATCCGGAAGTCTCCAGTTCACCTTCTGAGATTCCTCATCCGGCACACGAC 1579
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db 1580 CAGGAGGACCCAGTTCGACTTCACTTTCGTTGAAGTTTATCGAGTCAAGAAATTCACCTTC 1639
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
Db 1640 ACCTCGAAGCACCTGGAAATACGAGGACATGATCGAAGGAACAAGAGAGCAAGATTT 1699
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
Db 1700 GGAAGATCTGCAAGGACACACCTCTTGTCACTTGTCTCAGCTCCAGACCTCTGGAAC 1759
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 1760 TGGCATGGCAAGTCAATGACACAGCCCGGCTATTGAGTCAAGGTCCATCGCAGCTGGTG 1819
Qy 521 ArgLeuPheAlaArgGlyIleGluGlu-----AsnProLysProAspSer 535
Db 1820 CGCTCTTCTCGGGGAATCGAGGAAGAGTCAAGCAAGAAACCCCAAGCCCGGAGGC 1879
RESULT 3
BM479389
LOCUS
DEFINITION
AGENCOURT_6418742 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:550281
5', mRNA sequence.
```

BM479389 1063 bp mRNA linear EST 05-PEB-2002
AGENCOURT_6418742 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:550281
5', mRNA sequence.

```

ACCESSION   BM479389
VERSION     BM479389.1  GI:18528431
KEYWORDS    EST.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens (human)

REFERENCE   1 (bases 1 to 1063)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L14M12140 row: c column: 18
           High quality sequence stop: 665.

FEATURES             source
     1..1063
     /location=Qualifiers
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:5502281"
     /tissue_type="retinoblastoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_67"
     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
           Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
           Average insert size 1.75 kb. Library constructed by Life
           Technologies."

ORIGIN
Alignment Scores:
Pred. No.:      5,228-140      Length:      1063
Score:          1518.00      Matches:    301
Percent Similarity: 92.22%      Conservative: 7
Best Local Similarity: 90.12%      Mismatches: 19
Query Match:     52.56%      Indels:     7
DB:              4              Gaps:       3

US-10-631-958-10 (1-537) x BM479389 (1-1063)

Qy 155 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAen 174
Db 1 GCACCACTGTTCCCTTACCTTCCATCCACTGACATCATCGTTACTGAACATGCTAAT 60

Qy 175 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 194
Db 61 CAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTC 120

Qy 195 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSer 214
Db 121 GCGCGAGATGTTATGTTTCAGCGAGGTGTCGACCGGTCTGATTTGGGAGACGACAGGAGC 180

Qy 215 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 234
Db 181 GCGCGGGTGCAGCAACACACCCCGGGCTGTGCTGGTCCCGAGTAGCTCCGGATTGGA 240

Qy 235 IleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAla 254
Db 241 ATCATTCGCGAGGGTCAACGACTGCGTGTGTTACTCCACGTTGGGACACGACGACGCA 300

Qy 255 GluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerVal 274
Db 301 GAAACCTCGCGCTGCATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTGTCCTCAGTC 360

Qy 275 HisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGly 294

```

```

Db 361 CACCACAACGACACACTCTTCGCTACTCCGTTGCCCTGCTGGGTACGGCTTCTACGGG 420
Qy 295 AspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSer 314
Db 421 GACATCATCAAGGACAGTGAGAGAAACGGTGGTGGTCTTGGCAGATACGACTTTTCA 480
Qy 315 GlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAla 334
Db 481 GGTTTAAAGACCTTCTCTCCACACCTGCTATGAAGGGACAGTGTCTCTCTCCCTGCA 540
Qy 335 GlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCys 354
Db 541 CAACACACGGTGGGATCTCCAAGGATAGGAAGCCCTCCCGGACAGGATGCTTTGTC 600
Qy 355 ArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAla 374
Db 601 AGGCAAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGTGTTGGAAGCT 660
Qy 375 AlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThr 394
Db 661 GCGGAGGACGTGGAGGAGTGGCAAGTCTGTGTGGAAAGTTTCTGGCNCATCATGCCACA 720
Qy 395 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 414
Db 721 AACATGTCCTGTGTTGTGCGCGGAGCCCGANGGGCTCTCNCCTGCTGCCACTTNGA 780
Qy 415 AspGlySerSerAspLeuIleLeuLysCysSerArgPheAspPheLeuArgPhe 434
Db 781 GACGGTCTTCTGACCTCATCTCATCGGAAATGCTCCAGGNTCAATTTTCTGAGATTT 840
Qy 435 LeuIleArgHisThrAsnGlnAspGlnPheAspPheThrPhe---ValGluValTyr 453
Db 841 CTCTATCAGGACACCAACCCAGCAGGACAGTTTGTACTTCCCTTTTGGTTGAAGTTATC 900
Qy 454 ArgValLysLysPhe-GlnPheThrSerLysHisMetGluAsp-----GluAspSerAs 471
Db 901 CCCGTGAGAAAATCCCGAGTTTACTCTCAAAACCCCTTGGGAGGATGAGGACACGCA 960
Qy 471 pLeuLysGluGlyGly-----LysLysArgPheGly 481
Db 961 CTTCCAGGGGGGGGGGGGGGGGAAAAAAGCCCTTTTGGG 1000

RESULT 4
BQ054406
LOCUS      BQ054406
DEFINITION AGENCOURT 6771237 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5803668
5', mRNA sequence.
ACCESSION   BQ054406
VERSION     BQ054406.1  GI:19813746
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1059)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2040 row: a column: 13
High quality sequence stop: 601.
FEATURES             source
     1..1059
     /organism="Homo sapiens"
     /mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:5803668"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 5.63e-125 Length: 1059
Score: 1367.00 Matches: 276
Percent Similarity: 92.74% Conservative: 5
Best Local Similarity: 91.09% Mismatches: 14
Query Match: 47.33% Indels: 8
DB: 5 Gaps: 4

```

US-10-631-958-10 (1-537) x BQ054406 (1-1059)

```

Qy 133 LeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArg 152
Db 15 TTACTGGTATTATCAACCCGTTTGGAGAAAGACAGCAAGCGGATATATGAAGA 74
Qy 153 LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHis 172
Db 75 AAAGTGGCACCACCTGTTCCACCTTACCTCCATCACCACTGACATCATCGTTACTGAACAT 134
Qy 173 AlaAsnGlnAlaLysGluThrIleuTyrCluIleAsnIleAspLysTyrAspGlyIleVal 192
Db 135 GCTAATCAGGCCAACAGGAGACTCTGTATGAGATTAACTAGACAAATACGACGGCATCGTC 194
Qy 193 CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln 212
Db 195 TGTGTGGCGGAGATGTTATGTTTACGCGAGGTCTGCACGGTCTGATTGGAGACCGCAG 254
Qy 213 ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg 232
Db 255 AGGACGCGCGGGTCGACAGCAACACCCCGGGCTGTGCTGGTCCCGTAGCTCCGG 314
Qy 233 IleGlyIleIleProIleGlySerThrAspCysValCysTyrSerThrValGlyThrSer 252
Db 315 ATTGGAATCATTTCCGCGAGGTCAACGAGACTGCGGTGTGTACTCCACCGTGGGCACGAC 374
Qy 253 AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSer 272
Db 375 GACGAGAAACCTCGGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCC 434
Qy 273 SerValHisIleAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 292
Db 435 TCAGTCCACCAACAACAGCACACTCTTCGCTACTCCGTTCCCTGCTGGGCTACGGCTTC 494
Qy 293 TyrGlyAspIleIleLysAspSerGluLysIysArgTrpLeuGlyLeuAlaArgTyrAsp 312
Db 495 TACGGGGACATCATCAAGGACAGTGAGAGAAACCGGTGGTGGTCTTGGCAGATACGAC 554
Qy 313 PheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeu 332
Db 555 TTTTCAGGTTTAAAGACCTTCTCTCCACCACTGTATATGACCGACAGTGTCTCTTCTC 614
Qy 333 ProIleGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 352
Db 615 CCTGCACAACAACACCGTGGGATCTCCAAGGATAGGAAGCCCTCGCGGCGAGGATGCTTT 674
Qy 353 ValCysArgGlnSerLysGlnGlnLeuGluGlnLysAlaLeuTyrGlyLeu 372
Db 675 GTTTGCAGGCAAGCAAGCAGCTGAGGAGGAGCAGAAAGCAAGCACTGTATGGGTG 734

```

```

Qy 373 GluAlaAlaGluAspValGluGluTrpGlnValValCys-GlyLysPheLeuAlaIleAs 392
Db 735 GAAACTCGGAGGACGCTGGAGGAATGCAAGTCGTCTGTGGAAAGTCTCTGGCCATTAA 794
Qy 392 nAlaThr-AsnMetSerCysAla---CysArgArgSerProArgGlyLeuSerProAlaA 411
Db 795 TGCCACAAAACATGCCCTGTGCTTGTCCGCCCGGACCTCCAGGGCCCTCTCCCCCGG 854
Qy 411 la-----HisLeuGly---AspGlySerSerAspLeuIleLeuIle-----ArgLysC 426
Db 855 GCTGGCCCAACTTGGGAAGACGGGGTCTTCCGACCTCTCTTCTTCTTCTTCTGGGAAT 914
Qy 426 YsSer 427
Db 915 GCTCC 919

```

```

RESULT 5
BQ063738 1078 bp mRNA linear EST 02-APR-2002
LOCUS
DEFINITION AGENCOURT 6873251 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925382
5' mRNA sequence.
ACCESSION BQ063738
VERSION BQ063738.1 GI:19891754
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1078)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-@email.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Cloned and Sequenced by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2096 row: p column: 23
High quality sequence stop: 640.

```

FEATURES

```

source
1..1078
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5925382"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2.2e-120 Length: 1078
Score: 1321.00 Matches: 260
Percent Similarity: 92.31% Conservative: 4
Best Local Similarity: 90.91% Mismatches: 15
Query Match: 45.74% Indels: 7
DB: 5 Gaps: 4

```

US-10-631-958-10 (1-537) x BQ063738 (1-1078)

Qy 90 LysArgAlaArgArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGlu 109
Db 1 AAGAGACGACGACGCGCGTGGAGTGGCGCGAGGTGACTTTCTGGTGTCCAGAGGAG 60
Qy 110 GlnLeuCysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArg 129
Db 61 CAGCTGTGTCACTTGTGGCTCGAGACCCCTGCGGGAGAGTCTGGAGAGCTGACGTCCAGA 120
Qy 130 ProLysHisLeuValPheIleAenProPheGlyGlyLysGlyGlnGlyLysArgIle 149
Db 121 CCAAAAGCATTTACTGGTATTTATCAACCGTTTGGAGGAAAGGCAAGGCAAGCGGATA 180
Qy 150 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleVal 169
Db 181 TATGAAGAAGAAAGTGGCCACCTGTTCACCTTGTAGCTCCATCACCACATGACATCATCGTT 240
Qy 170 ThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAsp 189
Db 241 ACTGAACATGCTTAATCAGGCCAAGAGACTCTGTATGAGATTAAATAGACAAATACGCAC 300
Qy 190 GlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGly 209
Db 301 GGCATCGTCTGTGTGGCGGACATGTTATGTTTCAGCGAGGTGCTGCACGGTCTGATTGG 360
Qy 210 ArgThrGlnArgSerAlaGlyValAlaAspGlnAenHisProArgAlaValLeuValProSer 229
Db 361 AGGACGCGAGAGAGGCGCGGTGCACAGAACCCCGCGGCTGTGTGTCTGCCAGT 420
Qy 230 SerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrVal 249
Db 421 AGCTCCGGATTGGAAATATATCCCGAGGTCACGAGCTCGGTGTGTATCTCCACCGTG 480
Qy 250 GlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMet 269
Db 481 GGCACGACGACGCGAAGAACCTCGCGCTGCATATCGTTGTGGGACTCGCTGSCCATG 540
Qy 270 AspValSerSerValHisAenSerThrLeuLeuArgTyrTyrSerValSerLeuLeuGly 289
Db 541 GATGTGTCTCTAGTCACCAACACGACACTCTCTGCTACTCTGCTGTCTGCTGTGGGC 600
Qy 290 TyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysAspTrpLeuGlyLeuAla 309
Db 601 TACGGCTTCTACGGGACATCATCAAGACAGTGAAGAAACGGTGTGTGGTCTTGC 660
Qy 310 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrVal 329
Db 661 AGATAGGACTTTTCAGGTTTAAAGACTTCTCTCCACCACTGCTATGAAGGACAGGG 720
Qy 330 SerPheLeuProAlaGlnHisThrValGlySer---ProArgAsp-ArgLysProCysAr 348
Db 721 TCCTTCTCTCTGCAACACACGCGGTGGGATCTCCAGGGAATAAGAAAGCCCTGCC 780
Qy 348 gAla---GlyCys-----PheValCysArgGlnSerLysGlnGlnLeuGluGlu--- 364
Db 781 GGGCCAGGATGCTCTTTGGTTGGCCAGCCAGGCAAGCAAGCCAGCTTGAAGGACGG 840
Qy 365 ----GlnLysLysAla 368
Db 841 AGCCCAAGAAAGCC 856
RESULT 6
CK000755
LOCUS
DEFINITION AGENCOURT 16363467 NIH_MGC_221 Homo sapiens cDNA clone
IMAGE:30707875 5', mRNA sequence.
ACCESSION CK000755
VERSION CK000755.1 GI:38526789
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 797)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: NDAM1073 row: c column: 20
High quality sequence stop: 656.
FEATURES
source
1..797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:30707875"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_221"
/note="Organ: mixed; Vector: pYX-Asc; Site_1: EcoRI;
Site_2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNa size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. Average insert size
4-5Kb. Adaptors 5' (AATTCGCGACGAGG)3' and 5'd
(CTCTCGCGCG)3' 3' Linker sequence - GCGGCGCTGAGAGCC T18.
Sequencing primers 3'end: T3 promoter primer 5'd
(ATTACCCCTCACTAAAGGGA)3'. 5' End: T7 promoter primer 5'd
(TAATACGACTCACTATAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Note: this is a NIH_MGC
Library"
ORIGIN
Alignment Scores:
Pred. No.: 1.91e-118 Length: 797
Score: 1299.50 Matches: 250
Percent Similarity: 96.54% Conservative: 1
Best Local Similarity: 96.15% Mismatches: 8
Query Match: 45.00% Indels: 1
DB: 7 Gaps: 1
US-10-631-958-10 (1-537) x CK000755 (1-797)
Qy 152 ArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGlu 171
Db 9 CGGCACGAGGACCACTGTTCCACCTTAGCCCTCCATCACCCTGACATCATCGTTACTGAA 68
Qy 172 HisAlaAenGlnAlaLysGluThrLeuTyrGluIleAenIleAspLysTyrAspGlyIle 191
Db 69 CATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAATACGCGGCATC 128
Qy 192 ValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThr 211
Db 129 GTCTGTGTGCGCGGAGATGGTATGTTACGCGAGGTGTCACGGTCTGATTGGGAGGACG 188
Qy 212 GlnArgSerAlaGlyValAspGlnAenHisProArgAlaValLeuValProSerSerLeu 231
Db 189 CAGAGGACGCGCGGGTGCACAGAACCCCGGGCTGTGCTGGTCCCAGTAGCCCTC 248
Qy 232 ArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThr 251
Db 249 CGGATTGGATCATTTCCCGCAGGGTCAACGGACTGCGTGTGTACTCTCCACCGTGGGCACC 308

Qy 252 SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal 271
 Db 309 AGCGACGAGAAACCTCGGCGCTGCATATCTGTTGGGACTCGCTGGCCATGGATGTG 368
 Qy 272 SerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGly 291
 Db 369 TCCTCAGTCCACCAACACAGCACACTCTCTCGCTACTCCGTTGCTCCCTGGGTACGGC 428
 Qy 292 PheTyrGlyAspIleIleLysAspSerGluLysValysArgTrpLeuGlyLeuAlaAatGlyTyr 311
 Db 429 TTCTACGGGACATCATCAGGACAGTGAAGAAACGGTGGTGGGTCTTGCAGATAC 488
 Qy 312 AspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPhe 331
 Db 489 GACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTTC 548
 Qy 332 LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 351
 Db 549 CTCCTCGCACAAACACCGGTGGGATCTCCAAGGATAGGAAGCCCTCGCGGACGATGC 608
 Qy 352 PheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGly 371
 Db 609 TTTGTTTTCAGGCAAGCAAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGGT 668
 Qy 372 LeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIle 391
 Db 669 TTGGAAGCTCGGAGGACGTGGAGAGTGGNCAGTCTGTGTGGAAAGTTTCTGGGCCNA 728
 Qy 392 AsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAla 411
 Db 729 TCATGCAACAACATGCTCTGTGCTGTGCGCGAGGCCA---GGGGCTCTCCCGGTGCC 785

RESULT 7
 BQ057191
 LOCUS BQ057191 1047 bp mRNA linear EST 29-MAR-2002
 DEFINITION AGENCOURT_6769595 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5812382
 5', mRNA sequence.
 ACCESSION BQ057191
 VERSION BQ057191.1 GI:19816531
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1047)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Lou Staudt
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Cloning Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM2062 row: 1 column: 15
 High quality sequence stop: 535.
 Location/Qualifiers
 1. 1047
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5812382"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_99"
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory

ORIGIN
 Alignment Scores: Pred. No.: 1,02e-116 Length: 1047
 Score: 1284.00 Matches: 274
 Percent Similarity: 91.67% Conservative: 12
 Best Local Similarity: 87.82% Mismatches: 13
 Query Match: 44.46% Indels: 13
 DB: 5 Gaps: 2
 US-10-631-958-10 (1-537) x BQ057191 (1-1047)

Qy 133 LeuLeuValPheIleLeuAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArg 152
 Db 13 TTACTGGTATTTATCAACCCGTTTGGAGGAAAGGAACAAGCAAGCGGATATATGAAGA 72
 Qy 153 LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHis 172
 Db 73 AAAGTGGCACCACTGTTTACCTTAGCTCCATCACCCTGACATCATCGTTACTGAACAT 132
 Qy 173 AlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleVal 192
 Db 133 GCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGCGCATCGTC 192
 Qy 193 CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln 212
 Db 193 TGTGTGCGCGGAGATGGTATGTTTTCAGCGAGGTGTCGACGGTCTGATTTGGAGGACGAG 252
 Qy 213 ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg 232
 Db 253 AGAGAGCGCGGGTTCGACCAAGAACCCCGCGGGCTGTGTGGTCCCGACTAGCTCCGG 312
 Qy 233 IleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSer 252
 Db 313 ATTGGAATCATTTCCCGCAGGGTCAACGAGCTGCGTGTGTACTTCCACCGTGGGCACGAGC 372
 Qy 253 AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSer 272
 Db 373 GACGACAGAAACCTCGGCGCTGCATATGTTTGTGGGACTCGCTGGCCATGGATGTGTCC 432
 Qy 273 SerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 292
 Db 433 TCAGTCCACCAACACGACACACTCTTCGCTACTCCGTTGCTGGTGGGCTACGGGTTC 492
 Qy 293 TyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAsp 312
 Db 493 TACGGGGACATCATCAAGGACAGTGAAGAAACGGTGGTGTGGTCTTGGCAGATACGAC 552
 Qy 313 PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLe 332
 Db 553 TTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGGGACAGTGTCTTCTCT 612
 Qy 332 uProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 351
 Db 613 CCTGTGACAAACACACCGGTGGGATCTCCAAGGATAGGAACCTTGGCCGCGGACGATGC 672
 Qy 352 PheValCysArgGlnSerLysGlnGlnLeuGluGluGluGlnLysLysAlaLeuTyr 370
 Db 673 TTTGTTTTCGCGGCAAGCAAGCAGCAGCTGGAGGAGGAGGCCCAAGAAACCACTTGTAT 732
 Qy 371 GlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAl 390
 Db 733 GGTTTGGAACCTCGGGAAGACGGCAGAGTGGCAAAATCTTCTGGGGGGAGAGTTTCTGGC 792
 Qy 390 aIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSer---P 409
 Db 793 CATCCATGCCCAAAACCTGCTGGTGTGTCGCGGACCCCGGAGGGGCTTCTCCCC 852
 Qy 409 roAlaAlaHisLeuGlyAsp---GlySerSerAspLeuIleLeuIleArgLysCys-S 427

of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

Db 853 CGGCTGCCCAATTTGGGAGAGAGGTCTTCTGAACCTCTCTCTCAGGAAATGGCT 912
 QY 427 erArgPheAsnPhe 431
 Db 913 CCCGGTTCATTTT 926

RESULT 8
 CK000525
 LOCUS
 DEFINITION AGENCOURT 16369000 NIH MGC 221 Homo sapiens cDNA clone EST 26-NOV-2003
 IMAGE:30708597 5', mRNA sequence.

ACCESSION CK000525
 VERSION CK000525
 KEYWORDS EST.
 SOURCE CK000525.1 GI:38526559

ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 713)

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-remail.nih.gov
 Tissue Procurement: James Martin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

plate: NDAM1075 row: a column: 22
 High quality sequence stop: 689.

FEATURES

source
 1..713
 location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30708597"
 /lab_host="DH10B Tona"
 /clone_lib="NIH MGC 221"
 /note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;
 Site 2: NotI; Library is oligo-dT primed and directionally
 cloned. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated with
 EcoR I adaptor, digested with Not I and then cloned
 directionally into pYX-Asc vector. Average insert size
 4-5Kb. Adaptors 5'(AATCGGCACGAGG)3' and 5'd
 (CTCGTCCG)3'. 3' linker sequence - CGGCCGTGAGACC T18.
 Sequencing primers 3' end: T3 promoter primer 5'd
 (ATTACCCCTCACTAAAGGA)3', 5' End: T7 promoter primer 5'd
 (TAATACCACTCACTATAGG)3'. Library was constructed in the
 laboratory of M. Bento Soares. Note: this is a NIH_MGC
 Library"

ORIGIN

Alignment Scores:
 Pred. No.: 4,42e-116 Length: 713
 Score: 1275.00 Matches: 235
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 44.15% Indels: 0
 DB: 7 Gaps: 0

US-10-631-958-10 (1-537) x CK000525 (1-713)

QY 277 AnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 296

Db 9 AACAGCACATCCCTTCGCTACTCCGTGTCCTCTGGGCTACGGCTTCTACGGGACATC 68
 QY 297 IleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 316
 Db 69 ATCAAGGACAGTGAGAGAAACGGTGGTTGGTCTTCCAGATACGACTTTTCAGTTTA 128
 QY 317 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 336
 Db 129 AAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTCTCCCTGCAACAC 188
 QY 337 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 356
 Db 189 ACGTGGGATCTCCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTCAGGGCAA 248
 QY 357 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu 376
 Db 249 AGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGGTTTGGAGCTCGGAG 308
 QY 377 AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet 396
 Db 309 GACGTGGAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAACATG 368
 QY 397 SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 416
 Db 369 TCCTGTGCTTGTCCGGAGCCCGCAGGGGCTCTCCCGGCTGCCCATTTGGAGACGGG 428
 QY 417 SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle 436
 Db 429 TCTTCTGACCTCATCTCTCATCCGGAATGCTCCAGGTTCATTTTCTGAGATTCTCATC 488
 QY 437 ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys 456
 Db 489 AGGCACACCAACACGACGAGCAGCTTTGACTTCATCTTTTGTGAAGTTTATCGCGTCAAG 548
 QY 457 LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLeuGlyGly 476
 Db 549 AAATTCAGATTTCGTGAAGCAGCATGGAGATGAGGACAGCAGCCTCAAGAGGGGGGG 608
 QY 477 LysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsn 496
 Db 609 AAGAAGCGCTTTGGGCACATTTGCGAGCAGCCACCCCTCTCTGCTGTCACCGTCTCCAC 668
 QY 497 SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIle 511
 Db 669 AGCTCCTGGAACCTCGACGGGAGGTCTCTGCACAGTCTCTGCCATC 713

RESULT 9

CD655311

LOCUS

DEFINITION

AGENCOURT 14552675 KIA Human H1 Embryonic Stem Cell cDNA Library

(Long) Homo sapiens cDNA clone IMAGE:30426593 5', mRNA sequence.

ACCESSION

CD655311

VERSION

CD655311.1 GI:31895467

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 820)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-remail.nih.gov

Tissue Procurement: Irene Ginis

cDNA Library Preparation: Yulan Piao and Minoru Ko

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC c

lone distribution information

can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
 Plate: NDAM512 row: k column: 18
 High quality sequence stop: 673.

FEATURES source

1. 820
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="IMAGE:30426593"
 /tissue_type="Embryonic Stem cells"
 /cell_line="WA01"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
 /note="vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI;
 This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker IL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by extraction-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricion-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Alignment Scores:
 Pred. No.: 1.71e-114 Length: 820
 Score: 1260.00 Matches: 250
 Percent Similarity: 93.12% Conservative: 7
 Best Local Similarity: 90.58% Mismatches: 12
 Query Match: 43.63% Indels: 8
 DB: 6 Gaps: 5

US-10-631-958-10 (1-537) x CD655311 (1-820)

Qy 247 SerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSer 266
 Db 12 TCCACCGTGGCAGCAGCAGCAGAAACCTCGCGCTGCATATCGTTGTTGGGACTCG 71

Qy 267 LeuAlaMetAspValSerValHisAsnSerThrLeuLeuAgtTyrSerValSer 286
 Db 72 CTGGCCATGGATGTCTCAGTCACCAACAGACACACCTCTCGCTACTCTCGTGTC 131

Qy 287 LeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeu 306
 Db 132 CTGCTGGGCTACGGCTCTACGGGGACATCATCAAGGACAGTGAGAGAAACGGTGGTTG 191

Qy 307 GlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGlu 326
 Db 192 GGTCTTGCAGATACGACTTTTCAGGTTTAGAGACCTTCTCTCCACCACCTGCTATGAA 251

Qy 327 GlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 346
 Db 252 GGCACAGTGTCTCTCTCTCCCTGCACACACACCGGTGGGATCTCCAAGGATAGAGAGCCC 311

Qy 347 CysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLys 366
 Db 312 TGC CGGCGCAGGATGCTTTGTTGTCAGGCAAAAGCAGCAGCTGGAGGAGGAGCAAG 371

Qy 367 LysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGly 386
 Db 372 AAAGCAGCTGTATGGTGTGGAAAGCTGGGAGGAGCTGGAGGAGTGGCAAGTCGTCGTGGG 431

Qy 387 LysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly 406
 Db 432 AAGTTCTTGGCCATCAATGCCACACACATGCTCTGTGCTTGTGCGCGAGCCCCAGGGGC 491

Qy 407 LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeuLeuLeuLeuLysCys 426
 Db 492 CTCTCCCGCGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCATCCGGAATGC 551

Qy 427 SerArgPheAsnPheLeuArgPheLeuLeuArgHisThrAsnGlnGlnAspGlnPheAsp 446
 Db 552 TCCAGGTTCAATTTCTGAGATTTCATCAGGCACACCAACAGCAGGACCAAGTTTGAC 611

Qy 447 PheThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGlu 466
 Db 612 TTCACTTTCTTGAAGTTATCGGTCAAGAAATTCAGTTTAGTCTGAGCAGACATGGAG 671

Qy 467 AspGluAspSerAspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSer 486
 Db 672 GATGAGCAGCAGCACCTCAAGGAGGGGAGGAGCGCTTGGGCACAT---TGCAGCAGC 728

Qy 487 HisProSerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeu 506
 Db 729 ---ACCTCTGCTGCTGCAGC---TCTCAGAGCTCTG---GACTGCGAGCGGAGTCTG--- 777

Qy 507 HisSerProAla---IleGluValArgValHisCysGlnLeuValArg 521
 Db 778 -----CCAGCTGCATCAGTCAAGTCAATGCATGCATGCTGCATGCCAGA 819

RESULT 10
 CF135528
 LOCUS
 DEFINITION
 UI-HF-BNO-amf-g-10-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone
 IMAGE:3050211 5', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 732)
 Ronaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)

JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 97044477
 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Louis Staudt
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/humanfi.html


```

FEATURES
  source      Seq primer: pYX-5
              Location/Qualifiers
                1..732
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:3090211"
                /tissue_type="lymph"
                /cell_type="germinal center B cells"
                /cell_line="MGC85"
                /lab_host="DH10B (lTI)"
                /clone_lib="NIH MGC 50"
                /note="Vector: pYT3-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
ORIGIN
Alignment Scores:
  Pred. No.:      5.62e-113      Length:      732
  Score:          1244.00        Matches:    231
  Percent Similarity: 99.57%      Conservative: 0
  Best Local Similarity: 99.57%    Mismatches: 0
  Query Match:     43.07%        Indels:     1
  Db:              7            Gaps:         0
US-10-631-958-10 (1-537) x CF135528 (1-732)
QY 307 GlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGlu 326
DB 9  GGTCTTGCAGATAGACTTTTCAGGTTAAAGACTTCCTCTCCACCACTGCTATGAA 68
QY 327 GlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 346
DB 69 GGGACAGTGTCTTCTCTCCATGCACACACACGGTGGGATCTCCAAGGGATAGGAAGCCC 128
QY 347 CysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLys 366
DB 129 TCCCGCGGAGGATGCTTTGTTGCGAGGCAAGCAAGCAGCAGCTGGAGGAGGAGGAGAAG 188
QY 367 LysAlaLeuTyrGlyLeuGluAlaGluAspValGluGluTrpGlnValValCysGly 386
DB 189 AAAGACATGTATGTTTGAAGCTCGGAGGACGTGGAGGATGGCAAGTCTGCTGTGG 248
QY 387 LysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly 406
DB 249 AAGTTTCTGGCCATCAATGCCACAAACATGCTCTGTGCTTGTGCGCGAGCCCCAGGGGC 308
QY 407 LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCys 426
DB 309 CTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAATGC 368
QY 427 SerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAsp 446
DB 369 TCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAACAGCAGACACCATGTTGAC 428
QY 447 PheThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGlu 466
DB 429 TTCACCTTTGTTGAGTTTATCGCTCAGAAATTCAGTTTACGTCGAGGACATGAG 488
QY 467 AspGluAspSerAspLeuLys-GluGlyLysLysArgPheGlyHisIleCysSerSe 486
DB 489 GATGAGGACAGCGACCTCAAGGGAGGGGGAAGAAGCGCTTTGGGCACATTTGCAGCAG 548
QY 486 HisProSerCysCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValIle 506
DB 549 CCACCCCTCTGCTGTCACCGCTTCCCAACAGCTCTCTGGAACCTCGCAGCGGGAGTCT 608
QY 506 HisSerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGly 526
DB 609 GCACAGCCCTGTCATCGAGGTGAGGTCCACTGCCAGCTGCTGTTGCACTCTTTGCACGAG 668
QY 526 ylleGluGluAsnProLysProAspSerHisSer 537
DB 669 GATTGAAGAGAAATCCGAAGCCAGACTCACAGC 702
QY 526 ylleGluGluAsnProLysProAspSerHisSer 537
DB 669 GATTGAAGAGAAATCCGAAGCCAGACTCACAGC 702
RESULT 11
LOCUS CB246749
DEFINITION CB246749
ACCESSION CB246749
VERSION CB246749
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 758)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
              Location/Qualifiers
                1..758
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6"
                /db_xref="taxon:10090"
                /clone="IMAGE:6835595"
                /tissue_type="whole brain"
                /dev_stages="embryo 12.5dpc"
                /lab_host="DH10B (T1 phage resistant)"
                /clone_lib="NIH-BMAP_F10"
                /note="Organ: Brain; Vector: pYX-Asc; Site 1: Eco RI;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCCAGCAGC. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
ORIGIN
Alignment Scores:
  Pred. No.:      4.54e-109      Length:      758
  Score:          1205.00        Matches:    223
  Percent Similarity: 95.63%      Conservative: 18
  Best Local Similarity: 88.49%    Mismatches: 11
  Query Match:     41.72%        Indels:     0
  Db:              6            Gaps:         0
US-10-631-958-10 (1-537) x CB246749 (1-758)
QY 198 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 217
DB 1  GGCATGTTTCAGCAGAGGTCTGATGGGGTGATGGGAGGACGACAGCGCGCTGGTATC 60

```

```
Qy 218 AspGlnAenHisProArgAlaValLeuValProSerSerLeuArgLleGlyVilePro 237
Db 61 GACCCCATACCCCGGCGCGTGTGTGTCAGTACCTCAGATCGGATCATACCC 120
Qy 238 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 257
Db 121 GCAGGGTCCACAGATTGTGTGTACTCAACAGTGGGCACAAACGACGACAGACATCG 180
Qy 258 AlaLeuHisLleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsn 277
Db 181 GCATTGACATCATATTATTTGGGACTCACTGGCAATAGACGTGTCTCTGTGCATACCAT 240
Qy 278 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspLleLle 297
Db 241 AACACGCTGCTGCGGTACTCGGTTCCTGTCTGGGTACGGTTTCTACGGGGACTTATC 300
Qy 298 LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 317
Db 301 AAGGACAGTGAAGAAAGAAAGCGTGGCTCGTCCGGTATGATTTCTCAGGGTTGAAG 360
Qy 318 ThrPheLeuSerHisHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThr 337
Db 361 ACCTTTCTCTCATCTACTATGAGGGACACTGTCTCTCCAGCACACACACG 420
Qy 338 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 357
Db 421 GTGGATCTCCACGGGACAATAAACCTTCGCGGCTGGGTCTCGTGTGCGAGCAGC 480
Qy 358 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 377
Db 481 AAGCAACAGCTGGAAGAAAGAAAGAAAGCCCTGTATGCCCTGGGAACGCCGAGGAA 540
Qy 378 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaLleAsnAlaThrAsnMetSer 397
Db 541 ATGGAAGAGTGGCAAGTACATGTGGGAAGTTCCTGGCCATCAATGCCACCAACATGTCC 600
Qy 398 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 417
Db 601 TGTGCTTGTCTCGGAGCCCTGGGGCTGTCTGCCATTTGCCATCTGGGAGATGGGTCT 660
Qy 418 SerAspLleLleLeuLleArgLysCysSerArgPheAsnPheLeuArgPheLeuLleArg 437
Db 661 TCTGACCTCATCTTATCCGGAAGTCTCCAGGTTCANCTTCTCCTGAGATTCTCATCCGG 720
Qy 438 HisThrAsnGlnGlnAspGlnPheAspPheThrPhe 449
Db 721 CACAGAACACGAGGACCACTTCGACTCTCATTTC 756
```

```
RESULT 12
BX952302 LOCUS BX952302 653 bp mRNA linear EST 01-MAR-2004
DEFINITION DKFZp781L11183_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
            DKFZp781L11183_5', mRNA sequence.
ACCESSION BX952302
VERSION BX952302.1 GI:43428907
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 653)
AUTHORS Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
TITLE EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
```

consortium of the German Genome Project.

No s1 sequence available.

This clone (DKFZp781L11183) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

```
Location/Qualifiers
1..653
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781L11183"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/notes="Vector: pSport1_sfi; Site_1: sfiIA; Site_2: sfiIB;
cDNA-collection"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 8.84e-107 Length: 653
Score: 1181.00 Matches: 217
Percent Similarity: 100.00% Conservations: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.89% Indels: 0
DB: 5 Gaps: 0

US-10-631-958-10 (1-537) x BX952302 (1-653)
Qy 297 IleyAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 316
Db 2 ATCAAGGACAGTGAAGAAACCGTGTGGTCTTCCAGATACGACTTTTCAGGTTTA 61
Qy 317 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 336
Db 62 AAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTCGCACACAC 121
Qy 337 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 356
Db 122 ACGGTGGATCTCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGAGGCAA 181
Qy 357 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu 376
Db 182 AGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGACCTGTATGTTTGGAGCTGCGGAG 241
Qy 377 AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaLleAsnAlaThrAsnMet 396
Db 242 GACGTGAGAGAGTGGCAAGTCTCTGTGGGAGTTTCTGCCCATCAATGCCACAAACATG 301
Qy 397 SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 416
Db 302 TCCTGTGTTGTGCGCGAGGCCCGGAGGCTCTCCCGGCTGCCACTTGGGAGACGGG 361
Qy 417 SerSerAspLleLleLeuLleArgLysCysSerArgPheAsnPheLeuArgPheLle 436
Db 362 TCTTCTGACCTCATCTCTCATCCGAAATGCTTCCAGGTTCATTTTCTGAGATTCTCATC 421
Qy 437 ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys 456
Db 422 AGGCACACCAACACGACGAGGACAGTGTGACTTCTACTTTTGTGAAGTTATTCGGTCAAG 481
Qy 457 LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly 476
Db 482 AAATTCAGTTTACGTGGAAGCACATGAGGATGAGGACGACGACCTCAAGAGGGGGGG 541
Qy 477 LysLysArgPheGlyHisLleCysSerSerHisProSerCysCysThrValSerAsn 496
Db 542 AAGAAGCGCTTGGGACATTTGACGACGACCCCTCTCTGCTGCTGCTGCTCTCCAC 601
Qy 497 SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaLleGluVal 513
Db 602 AGCTCTGGAACTCGGAGGGGAGGTCTTGACAGAGCCCTGCACTCGAGGTC 652
```

RESULT 13

CN296312
 LOCUS 17000532197874 GRN_EB Homo sapiens cDNA 5', mRNA linear EST 16-MAY-2004
 ACCESSION CN296312
 VERSION CN296312.1 GI:47312726
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 661)
 Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
 Transcriptional characterization elucidates signaling networks that control human ES cell growth and differentiation
 Nat. Biotechnol. 22 (6), 707-716 (2004)
 Contact: Brandenberger R
 Regenerative Medicine
 Geron Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@geron.com
 Insert Length: 661 Std Error: 0.00.

FEATURES
 source
 1..661
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, embryoid bodies
 derived from H1, H7 and H9 cells"
 /clone_lib="GRN_EB"
 /note="oligo dt primed, full-length enriched cDNA library
 from embryoid body outgrowths derived from hES cell lines
 H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
 conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 1,12e-105 Length: 661
 Score: 1170.00 Matches: 219
 Percent Similarity: 99.55% Conservative: 0
 Best Local Similarity: 99.55% Mismatches: 0
 Query Match: 40.51% Indels: 1
 DB: 7 Gaps: 0

US-10-631-958-10 (1-537) x CN296312 (1-661)

Qy 278 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIlelle 297
 Db 3 AGCACACTCTCGCTACTCCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATC 62

Qy 298 LysAspSerGluLysLeuArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 317
 Db 63 AAGGACGTGAGAGAAACGGTGGTGTGGCTCTGCCAGATACGACTTTCAGGTTTAAAG 122

Qy 318 ThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 337
 Db 123 ACCTTCCTCTCCACCATCTGTATGAGGGACAGTGTCTTCTCCCTGCGACACACAG 182

Qy 338 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 357
 Db 183 GTGGGATCTCCAAGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTTTCAGGCAAGC 242

Qy 358 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 377
 Db 243 AAGCAGCAGCTGGAGGAGGAGCAGAGAAAGACACTGTATGTTTGGAAAGCTCGCGAGGAC 302

Qy 378 ValGluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 397
 Db 303 GTGGAGGAGTGGCAAGTCTGTCTGGGAAGTTTCTGGCCATCATATGCCACCAACATGTCC 362

Qy 398 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 417
 Db 363 TGTGCTTGTGCCGGAGCCCGCAGGGGCTCTCCCGGGCTGCGGAGAGCGGGTCT 422

Qy 418 SerAspLeuLeuLeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuArg 437
 Db 423 TCTGACCTCATCTCATCCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG 482

Qy 438 HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 457
 Db 483 CACACCAACAGCAGGACCACTTTGACTTCACTTTTGTGGAAGTTTATCGCGTCAAGAAA 542

Qy 458 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGlu-GlyGlyLys 477
 Db 543 TTCCAGTTTACGTGGAAGACATGAGGATGAGACAGCGACCTCAAGAGGGGGGGAA 602

Qy 477 sLysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsn 496
 Db 603 GAAGCGCTTTGGGACACATTTGCAGACGCCCTCTCTGCTGCTGCACCGTCTCCAC 660

RESULT 14
 CK603033
 LOCUS
 DEFINITION AGENCOURT 17899852 NIH_MGC_234 Rattus norvegicus cDNA clone
 IMAGE:7193195 5', mRNA sequence.
 ACCESSION CK603033
 VERSION CK603033.1 GI:41116352
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 812)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@ps-remail.nih.gov
 Tissue Procurement: Howard Jacobs
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LIAM15055 row: b column: 09
 High quality sequence stop: 657.
 Location/Qualifiers
 1..812
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7193195"
 /tissue_type="heart, pooled"
 /lab_host="DH10B Tona"
 /clone_lib="NIH_MGC_234"
 /note="Organ: heart; Vector: pExpress-1; Site:1: EcoRV;
 Site:2: NotI; RNA obtained from pooled heart tissue from a
 mix of male and female animals at 8 wk old. Tissues were
 snap-frozen and kept at -80C for two days before RNA
 extraction and purification (Tri-reagent method). cDNA was
 primed using oligo-dT primer:
 5'-pGATAGTTCATGACGCGCGCCGCCCTT25-3' and cloned into
 the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb
 resulted in an average insert size of 2.2 kb. This primary
 library is normalized (non-normalized primary library is
 NIH_MGC_233) and was constructed by Express Genomics
 (Frederick, MD). Note: this is a NIH_MGC library."

ORIGIN

```
Alignment Scores:
Pred. No.:      6.14e-94      Length:      812
Score:          1053.50      Matches:    206
Percent Similarity: 85.23%      Conservative: 19
Best Local Similarity: 78.03%      Mismatches: 35
Query Match:    36.48%      Indels:    5
DB:              7          Gaps:      3

US-10-631-958-10 (1-537) x CK603033 (1-812)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20
   |||||
Db 15 ATGGCGCAATGGGGCGGCGAGCCGCTGCAATCGTGTGCTGGTGAAGCAGCAACGC 74

Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
   |||||
Db 75 TGTCCCGTCAGCTTGGAGCCGCGCGGCTCTGCTAGCTGGTGGCGAGCCCGAGCC 134

Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCyaserValProValSerGluIleAlaVal 60
   |||||
Db 135 GGGCCCTCGGCCCGCGATCCGATGCTTACTCGGTACAGTGTGAGATCATCGCCGTT 194

Qy 61 GluGluThrAspValHisGlyHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
   |||||
Db 195 GAGGAAAAGAGTCCACCAGAAACAGTCTCCCAATGGCCGATGCATAGAAATGGA 254

Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100
   |||||
Db 255 CCGTTTGCATTCACAGTCCACTGTGTGAAGCGAGCTCGACACCCGCTGGAAGTGG 314

Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
   |||||
Db 315 CCGGTGACCTTCTGGAGCCCGCAGCAGCAGTGTCTCACTGTGCTGCAGACCTCG 374

Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
   |||||
Db 375 GAGCTGCTGGAGAACCTGACTTCAAGACCGAAGCATTTGCTGTATTATCAACCC 434

Qy 141 GlyClyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
   |||||
Db 435 GGAGGAAAGGCCCGAGCCAGCACATCTATGAAAAGAAAGTGGCACCTTTGTTCC 494

Qy 161 AlaSerIleThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
   |||||
Db 495 GCTTCCATCACCATGAGATCATCATTTACTGAGCATGCCAATCAAGCCAGGAGACT 554

Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
   |||||
Db 555 TACGAGATCAACACAGACAGCTATGACGGCATCTGTGTGTGTTGGTGGGCGGCATG 614

Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
   |||||
Db 615 AGCGAGGTGCTGCACGGGGTAATTTGGGAAGACGACGAGCAGCGCTGGTGTGTC 674

Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
   |||||
Db 675 CACCCCGAGCCGTGCTGGTCCAGTACCCT--AGGATCGGCATCATCTCGGCGGTAC 732

Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
   |||||
Db 733 AGA--TGTGTGTGC---TCTCACCGTGGCACAACG-----AGCAGAAACGTGCTTACC 780

Qy 261 IleValValGly 264
   |||||
Db 781 ATCATATTGGGA 792

RESULT 15
BP224560
LOCUS
DEFINITION
BP224560 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA
clone DAT01067, mRNA sequence.
ACCESSION
BP224560

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.:      4.96e-90      Length:      581
Score:          1012.00      Matches:    192
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.48%      Mismatches: 0
Query Match:    35.04%      Indels:    0
DB:              5          Gaps:      0

US-10-631-958-10 (1-537) x BP224560 (1-581)

Qy 184 AsnIleAspLysTyrAspGlyIleValCysValGlyAspGlyMetPheSerGluVal 203
   |||||
Db 3 AACATAGACAAATACACGCGCATCGTCTGTGTCGGCGGAGATGATGTTTCAGGAGGTG 62

Qy 204 LeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArg 223
   |||||
Db 63 CTGCACGGTCTGATTGGGAGGACGACAGAGGCGCGGGCTCGACGAGGACCCCGG 122

Qy 224 AlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCys 243
   |||||
Db 123 GCTGTGTGCTGCCAGTAGCCTCCGGATTGGAATCATTTCCCGCAGGGTCAACGAGCTGC 182

Qy 244 ValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValVal 263
   |||||
Db 183 GTGTGTTTACTCCACCGTGGGCACGACGACGAGAAACCTCGGCGCTGCATATCGTTGTT 242

Qy 264 GlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyr 283
   |||||
Db 243 GGGGACTCGTGGCCATGGATGTCTCTCAGTCCACACACAGCAGCAGCTCTTCGCTAC 302

Qy 284 SerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLys 303
   |||||
Db 303 TCCGTGTCCCTGTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGAAGAGAA 362

Qy 304 ArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHis 323
   |||||
Db 363 CGGTGGTGGGTCTTGCACATAGACTTTTTCAGGTTTAAAGACTTCTCTCCACAC 422

Qy 324 CysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAsp 343
   |||||
Db 423 TGCTATGAAGGGACAGTGTCTTCTCTCCCTGCACACACACGCGTGGGATCTCCAAAGGAT 482

Qy 344 ArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGluGlu 363
   |||||
```

Db 483 AGGAAGCCTGCCGGCAGGATGCTTTGTTTCAGGCAAGCAAGCAGCTGGAGGAG 542

Qy 364 GluGlnLysLysAlaLeuTyrGlyLeuGluAlaGlu 376

Db 543 GAGCAGAAAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAG 581

Search completed: September 6, 2005, 20:29:36
Job time : 5188.87 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:10:49 ; Search time 23.7411 Seconds
(without alignments)
2176.332 Million cell updates/sec

Title: US-10-631-958-10
Perfect score: 2888
Sequence: 1 MGATGAEPLOSVLWVKQR.....QLVRLFARGIEENPKPDHS 537

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	567.5	19.7	549	2 T33517	hypothetical prote
2	417.5	14.5	1240	2 T05162	hypothetical prote
3	291.5	10.1	458	2 T38776	hypothetical prote
4	248.5	8.6	687	2 S51398	hypothetical prote
5	245.5	8.5	473	2 T19707	hypothetical prote
6	210.5	7.3	624	2 S67059	hypothetical prote
7	189.5	6.6	310	2 A61665	hypothetical prote
8	176	6.1	310	2 A11293	hypothetical prote
9	147	5.1	303	2 F69795	conserved hypotet
10	136.5	4.7	306	2 A11769	conserved hypotet
11	136	4.7	732	2 T16422	hypothetical prote
12	130.5	4.5	309	2 A11528	conserved hypotet
13	129	4.5	295	2 A83894	hypothetical prote
14	126.5	4.4	306	2 A81394	conserved hypotet
15	125.5	4.3	295	2 D83734	hypothetical prote
16	125.5	4.3	309	2 A11171	conserved hypotet
17	122.5	4.2	433	2 S75948	hypothetical prote
18	115.5	4.0	294	2 G95120	conserved hypotet
19	114	3.9	364	2 F84898	hypothetical prote
20	112	3.9	345	2 E69678	involved in polyke
21	111.5	3.9	311	2 C97990	conserved hypotet
22	105	3.6	333	2 F71006	hypothetical prote
23	103.5	3.6	1028	2 A96719	hypothetical prote
24	102.5	3.5	315	2 A89978	conserved hypotet
25	102	3.5	297	2 F69595	multidrug resistan
26	102	3.5	343	2 JC7183	cathepsin Q (EC 3.
27	100	3.5	650	1 JCI450	fibroblast growth
28	99	3.4	1555	2 T18688	hypothetical prote
29	99	3.4	1973	2 G89608	protein B0272.5 [1

RESULT 1

T33517
hypothetical protein T10B11.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33517
R:Minx, P.; Kemp, K.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid T10B11.
A:Reference number: Z21363
A:Accession: T33517
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-549 <MIN>
A:Cross-references: UNIPROT:Q9TZ11; EMBL:AF098993; PIDN:AAC67466.1; GSPDB:GN00019; CESP:
A:Experimental source: strain Bristol N2; clone T10B11
C:Genetics:
A:Gene: CESP:T10B11.2
A:Map position: 1
A:Introns: 26/1; 76/2; 109/3; 159/1; 187/3; 229/3; 353/1; 398/2; 434/1; 467/3

Query Match	19.7%	Score	567.5	DB	2	Length	549
Best Local Similarity	30.8%	Pred. No.	3.4e-39				
Matches	139	Conservative	91	Mismatches	160	Indels	61
Gaps	13						
Qy	89	VKRARRHRKWAQ--VTFWCPEQLCHLWLTQLREMLEKLTSPKHLVFNPFQGGKGG	146				
Db	123	VYKDKQKWLKQIPVIFYTTSR--DYWHSLIOTTLRRVKNRPANIIIFINPFGNGKA	180				
Qy	147	KRIYERKVAFLTLA-SITTDIIIVTEHANOAKETLYEINIDKY---DGIVCVGGDGMFSE	202				
Db	181	QKIFKDNVDPAFFWLTPLGLRYKVLTERRANHARDYIVEMPPQWSAIDGLSVSGDGLFNE	240				
Qy	203	VLHGLIGRTQBSAGVDQNHPRAVLPVPSLSRIGIIPAGSTDCVCTVGTSTDAETSAHLIV	262				
Db	241	LLSGALLRTQDAGRINDPSSHLVTPHIREGIIAGSANSIVSTVHETNDHATSAVHIA	300				
Qy	263	VGDSLAMDVSSVHNSLTLLRYSVSLGFGVFGYIIDKSEKKRWGLARYDPSGLKTFLSH	322				
Db	301	IGSECNVDVCTVHOHQKLIRISANAISYMGVLDGVLDRSEYRCLGPIRYQMSALRTTIRH	360				
Qy	323	HCYEGTSVFTPAQHTVGSRRDR-KPCRAGCFVCRQSKQQLFEQEKALYGLEAAEDVEEW	381				
Db	361	PIYKGMVQFSLSHKENVNPDKQLPCLPCPCVMK-----PQNDKDYDHHWA----EFT	411				
Qy	382	QVVGKFLAINATNMNSCACRRSPRGLSPAHLGGSSDLILIRKCSRFNFRILIRHT--	439				
Db	412	HVICCVIPTVTPF-----TPYGLAFTGIGDGTLDLALVPRISRFRHNNQFMKRVAMY	463				
Qy	440	--NQDQDFTFVEVRYVKKFQFTSKHMEDESDLKEGKKRFGHICSSHSPSCCCTVSN	497				
Db	464	GGKQLYELD-PSLNCYRVTKWSY-----QPDADQEDFG-----	495				

hypothetical prote
protein C23F12.1 [
hypothetical prote
probable membrane
conserved hypotet
hypothetical prote
conserved hypotet
hypothetical prote
conserved hypotet
indoleacetamide hy
probable ClpA/B pr
hypothetical prote
ornithine decarboxyl
ubiquitin carboxyl
hypothetical prote


```
Qy 103 TFWCEQOLCHLWLTLEMLEKL-----TSRPKHLVFINPFGGKGQKRIYERKA 155
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 SLYIDYKPHSSHLKEEDDLVEEILKRYSQNTRRNKSIFVINPFGGKGAKCLFMTRAK 293
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 156 PLFTLASITTDIIVTEHANOAKETLYEINIDKYGIVCGDGMFSEVLHGLIGHTORSA 215
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 PLLLASCSIEVVYTKYCHAEIAREMDIDKYDTIACASGDGIPHEVINGLYQR----- 348
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 216 GVDQNHPRAVLPVSSLRIGIIPAGSTDCVYSTVGTSDAETSALHIVVGDSLAMDVSVH 275
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 -----PDHVRAFNNIAITEIPCGSGNAMSVSCHWNNPSYSTLCLIKSIETRIDLMCCS 402
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 276 HNSTLLRY-SVSLIG--YGYGDIIDKSEKKRWGLARYDF----- 313
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 403 QPSYAREHPKLSFSLQTYGLIAETDINTEFIRWMGPARFELGVAFNIIQKKYCEIYVK 462
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 314 -----SLGKTLFSLHHCYEGTVSPLPAQHTVGSPRDRKPCRAGCF----- 352
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 463 YAAKSKNELKHYLEHKKGSLEF---QH-ITMKNKDNDCCNYYENYETENEDDEDA 518
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 353 -----VCRSQKQLEEQKALYGLEAAEDV-BEQ-----VVCG 386
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 519 DADDEDHLISRDLADSSADQIKEDDFKIKYPLD--EGIPSDWERLDPNISNLLGIFVTG 576
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 387 KFLAINATMNSCACRRSPRGLSPAHLGDGSSDLIL 422
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 577 KMPYVAADTK-----PPPAALPSDGTMDMWI 602
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
T19707
hypothetical protein C34C6.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
R:Percy, C.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z19167
A:Accession: T19707
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-473 <W1>
A:Cross-references: UNIPROT:Q18425; EMBL:Z66494; PIDN:CAA91259.1; GSPDB:GN000020; CESP:C3
A:Experimental source: clone C34C6
C:Genetics:
A:Gene: CESP:C34C6.5
A:Map position: 2
A:Introns: 82/1; 126/1; 158/3; 276/1; 311/3; 427/3

Query Match 8.5%; Score 245.5; DB 2; Length 473;
Best Local Similarity 24.3%; Pred. No. 2e-12;
Matches 109; Conservative 66; Mismatches 159; Indels 115; Gaps 18;

Qy 124 EKLTS-----RPK-----HLVFINPFGGKGQKRIYERKVAFLFLASITTDIIVTE 171
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 EQLTSVILSRKPPPOEQCRGNLLVEINPNSGTGKSLFTFANTVGPGLDKSLIRYEVVVTT 126
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 172 HANOAKETLY-EINIDKYGIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSS 230
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 GPNHARVNLMTKADLGKNGVLLSGDLGVFEALNGILCRE-----AFRIPT 175
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 231 LRIGIIPAGSTD---CVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSL 287
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 176 LPIGIVPSGSGNGLLCSVLSKYGTQWNEKS----VMERALEIATSPTAKAESVALYSVK 231
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 288 -----LCYGYGDIIDKSEK-KRWGLARYDFSLGKTLFSLHHCYEGTVSFLPAQH 336
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 232 DNQSYASFLSIGWGLMADIDIDSEKWRKSLGHHRTVMGFIKSCNLSYKGLRYRYPK- 290
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 337 TVGSPDRDKPCRAGCFVCRSQKQLEEQKALYG----- 371
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 291 ----PKGFHPSNVFSVYEKTTQQRIDDSKVKTNGSVSDSEETMETKFNQWTLPDSD 346
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Qy 372 -LEAAEDVEEHQVCGKFLAINATMNSCACRRSPRGLSPAHLGDGSSDL--ILIRK-S 427
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 347 LAVGSSDLEETVVDNEDFNFIYAVTLSHIAADGP--FAPSAKLEDNRHLSYILWKDITG 404
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 428 RNFNLFRLIRHTNOODQDFDTFVEYVRVKFQFTSKMEDESDLK---EGGKKRFGHIC 484
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 405 RVNIAKYLLA-IEHETHLDLPFV-----KHVEVSSMKLEIVISEG-----SHV 446
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 485 SSHPSCCCTVSNSSWNCDEVLHSPAIEV 513
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 447 -----LDGEVVDTKTIEV 459
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
S67059
hypothetical protein YOR171c - Yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O3615
C:Species: Saccharomyces cerevisiae
C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
R:Bordonne, R.; Camaseses, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarassov, I.A.; Winsor
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67032
A:Accession: S67059
A:Molecule type: DNA
A:Residues: 1-624 <BOR>
A:Cross-references: UNIPROT:Q12246; EMBL:Z75078; NID:g1420415; PID:g1420417; GSPDB:GN000
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:LCB4; MIPS:YOR171C
A:Cross-references: SGD:S0005697
A:Map position: 15R

Query Match 7.3%; Score 210.5; DB 2; Length 624;
Best Local Similarity 26.3%; Pred. No. 2.4e-09;
Matches 66; Conservative 46; Mismatches 102; Indels 37; Gaps 6;

Qy 118 TLRMLEKL---TSRPKHLVFINPFGGKGQKRIYERKVAFLFLASITTDIIVTEHAN 174
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 211 TVBEILEKSYENSKRNSILVINPHGKGKTAKNLFITKARPILVESGCKIEIAYTKVAR 270
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 175 QAKETLYEINIDKYGIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIG 234
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 271 HAIDIAKDLDISKYDTIACASGDGIPYEIVINGLYRRPDR---VDAFNKLAVTQ----- 320
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 235 IIPAGSTDVCYSTVGTSDAETSALHIVVGDSLAMDV---SSVHHNSTLLRYSVSLG 291
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 321 -LPCGSGNAMSISCHWNNPSYALCLVKSIETRIDLMCCSQPSYMNEMWPRLSFLSQT 379
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 292 FYGDIIDKSEKKRWGLARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPPDRKPCRAGC 351
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 VIAESDINTEFIRWMGPVRFNLG----VAFNIIQG-----KKYPCVEFV 419
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 352 FVCRSQKQLE 362
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 KYAAKSKKELK 430
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
AG1665
hypothetical protein homolog lin1865 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG1665
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Fshih, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
```

A;Accession: AG1665
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-310 <GLA>
A;Cross-references: UNIPROT:Q8Y2AQS; GB:AL592022; PIDN:CAC97095.1; PID:gl6414366; GSPDB:G
A;Experimental source: strain Clip1262
C;Genetics:
A;Gene: lin1865

Query Match 6.8%; Score 189.5; DB 2; Length 310;
Best Local Similarity 21.8%; Pred. No. 5.3e-08;
Matches 91; Conservative 57; Mismatches 139; Indels 131; Gaps 16;

QY 131 KHLIVFNPFGKGQKRIYERKVAFLFTLAS-----ITTDIIVTEHANOAKETLYEINID 186
DB 3 KHARVIYNP-----TSGREIIKKNLADVLSILEQAGYVTSAHATTAEFGDAKHAEEAVRD 58

QY 187 KYDGI VCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVCY 246
DB 59 RYDLVVAAGDGTINEVINGIAEKYRP-----KVGIPTGTNDPFR 101

QY 247 STVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGVGYFGYDIIKO--SEKKR 304
DB 102 ALHVPDRVIVKATKIIAAGQSVAMDIGKA--NDT---YFINIGGGRLTEITYDVPSRLKT 156

QY 305 WLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSKQQLLEE 364
DB 157 MLGQLAYLLKG-----IEMLPs-----LKATKVKVEYD 184

QY 365 QKXALYGLEAAEDVEEVQVCGKFLAINATNMSCACRRSPRGLSPAHLGGSSDLILIR 424
DB 185 Q-----GVFEGEVM---FLLGLTNSIGGFEK---IAPDAKLDDGKFLSLIIVK 226

QY 425 KCSRNFNLRFL 435
DB 227 KVNLAEFIRLV 237

RESULT 8
A11293
hypothetical protein lmol1753 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: A11293
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: A11293
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-310 <GLA>
A;Cross-references: UNIPROT:Q8Y6D4; GB:NC_003210; PIDN:CAC99831.1; PID:gl6411207; GSPDB:
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmol1753

Query Match 6.1%; Score 176; DB 2; Length 310;
Best Local Similarity 22.2%; Pred. No. 7.1e-07;
Matches 69; Conservative 51; Mismatches 109; Indels 82; Gaps 11;

QY 131 KHLIVFNPFGKGQKRIYERKVAFLFTLAS-----ITTDIIVTEHANOAKETLYEINID 186
DB 3 KHARVIYNP-----TSGREIIKKNLADVLSILEQAGYVTSAHATTAEFGDAKHAEEAVRN 58

QY 187 KYDGI VCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVCY 246
DB 59 RYDLVVAAGDGTINEVINGIAEKYRP-----KVGIPTGTNDPFR 101

QY 247 STVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGVGYFGYDIIKO--SEKKR 304
DB 102 ALHVPDRVIVKATKIIAAGQSVAMDIGKA--NET---YFINIGGGRLTEITYDVPSRLKT 156

QY 305 WLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSKQQLLEE 364
DB 157 MLGQLAYLLKG-----IEMLPs-----LKATKVKVEYD 184

QY 365 QKXALYGLEAAEDVEEVQVCGKFLAINATNMSCACRRSPRGLSPAHLGGSSDLILIR 424
DB 185 Q-----GVFEGEVM---FLLGLTNSIGGFEK---IAPDAKLDDGKFLSLIIVK 226

QY 425 KCSRNFNLRFL 435
DB 227 KVNLAEFIRLV 237

RESULT 9
F69795
conserved hypothetical protein yerQ - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: F69795
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidos, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winter, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69795
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-303 <KUN>
A;Cross-references: UNIPROT:O31502; GB:Z99107; GB:AL009126; NID:g2632866; PIDN:CAB12492.1
A;Experimental source: strain 168
C;Genetics:
A;Gene: yerQ

Query Match 5.1%; Score 147; DB 2; Length 303;
Best Local Similarity 21.7%; Pred. No. 0.00018;
Matches 73; Conservative 54; Mismatches 120; Indels 90; Gaps 15;

QY 131 KHLIVFNPFGKGQKRIYERKVAFLFTLASITTDIIVTEHANOAKETLYEINID 187
DB 2 KHARVIYNP-----TSGREIFKHLAQVLQKFEQAGYVTSHTATTCAGDATHAAKEAALRE 57

QY 188 YDGI VCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVCY 246
DB 58 FDLIIAAGDGTINEVNVGL-----APLD-NRP-----TLGVIPTGTNDPFR 100

QY 247 STVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGVGYFGYDIIKO--SEKKR 304
DB 101 LGIPREDILKAADTVINGVARDIDIGQVNG-----QFINIAGGRLTEITYDVPSRLKT 155

QY 305 WLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSKQQLLEE 364
DB 156 MLGQLAYLLKGME-----MLPS-----LRPTEVEIYD 183

QY 365 QKXALYGLEAAEDVEEVQVCGKFLAINATNMSCACRRSPRGLSPAHLGGSSDLILIR 424

Db 184 GK--LFGGEI-----MLFLVLTUNSVGGFEK-----LAPDSSLNDGMFDMILK 225
QY 425 KCSRNFNLFRLI-----RHTNQDQDFTEFVEVYRVK 456
Db 226 KANLAEFIRVATMALRGEHINDQ---HIITYTKANRVK 259

RESULT 10
AH1769
conserved hypothetical protein lin2702 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH1769
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kunst, M.; Kuhn, M.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1769
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-306 <GLA>
A:Cross-references: UNIPROT:Q927T6; GB:AL592022; PIDN:CAC97928.1; PID:g16415238; GSPDB:C
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2702

Query Match 4.7%; Score 136.5; DB 2; Length 306;
Best Local Similarity 19.6%; Pred. No. 0.0014;
Matches 62; Conservative 55; Mismatches 91; Indels 109; Gaps 14;

QY 131 KHLVFINPFGKGQKRI-----YERKVAPLFTLASITTDIIIVTEHANOAKE 178
Db 3 KXAMIIYNPAGKNFRKLLPDAEKILTNADFEVLTP-STPAPKSTTLI-----AKQAAE 57

QY 179 TLYEINIDKIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPVPSLSRIGIIPA 238
Db 58 AGYEV-----VIAAGDGTVEVNGLMQVEKP-----KLGIILPV 93

QY 239 GSTDCVCVSTGTSDAETSALHIVVGDLSLMDVSVVHNSHTLLRYS-----VSLIGYGFYG 294
Db 94 GTTNDYARALNPAKDPLEALQIIAQOETIRVDIGKANETEPFINNAAGGRITEITYA--- 150

QY 295 DIIDSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVC 354
Db 151 --VKESMSKW-GRLAYLFSGL-----TVLP----- 173

QY 355 RQSKQOLEEBEQKALYGLEAAEDVEWVCGKFLAINATN-----MSCACRRSPRGLSPA 410
Db 174 -----KLSPVNVVEIYN-----EEIFKGELI---LFFVNKTVSGVMETLC-----pp 213

QY 411 AHLGDGSDLLILIRKCS 427
Db 214 AQLNSGMPPELLILKKVS 230

RESULT 11
T16422
hypothetical protein F52C9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16422
R:Ravello, T.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F52C9.
A:Reference number: Z18511
A:Accession: T16422
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-732 <FAV>
A:Cross-references: EMBL:U39850; NID:g1055052; PID:g1055058; PIDN:AAA81060.1; CESP:F52C9C
C:Genetics:
A:Gene: CESP:F52C9.3
A:Introns: 63/1; 106/3; 148/2; 303/1; 347/1; 408/2; 650/2

Query Match 4.7%; Score 136; DB 2; Length 732;
Best Local Similarity 19.9%; Pred. No. 0.005;
Matches 76; Conservative 56; Mismatches 128; Indels 122; Gaps 18;

QY 128 SRPKHLVFINPFGKGQKRIYERKVAPLFTLASITTDIIIVTEHANOAKETLYEINIDK 187
Db 65 TRPKRVFLVNVVNGSNGCCFDQFNKNALPLFLHLAGQVDVVVKADNQALEALAGAVDTQE 124

QY 188 YDGIVCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPVPSLSRIGIIPAG----- 239
Db 125 ADILVWVGDDTIGTVVTGIFRNREK-----AQLPVGFYPGGYDNLWLKR 169

QY 240 -----STDCV---CYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSHTLL-RYSVSL 288
Db 170 MLPVSFVNSDDVRHACETAMAVIEDQKSVY-----AFELTT--EGSTLAPEYGLGDV 220

QY 289 GYGFVGDIIKDSEKKRW-LGLAR---YDFSGLK---TFLSHHCYEGTVSFLPAQHTVG 339
Db 221 SAGHFRQ-IEDTRKKFWFSMAKRWYFWMLKRGAPIECHVEYET----- 268

QY 340 SPRDRKPCRCAGCFVCROSKQOLEEBEQKALYGLEAAEDVEWVCGKFLAINATNMSCA 399
Db 269 -----C-AGCEKCR-PKPIIEAPQWR-----WWHVLTG----- 294

QY 400 CRRSPRGLSPAHLGDGSSDL--ILIRKCSRNFEL-----RFLIRHTNQDQDFTEFVEV 452
Db 295 -----TPRYKNNDGQKDYTGIIINEKCGEKHELDTHGAEFLIENEQMSD----- 337

QY 453 YRVKKFOFTSKHMEDESDLKE 474
Db 338 YSQIRFRMGDPYMPPEEFWNE 359

RESULT 12
AH1528
conserved hypothetical protein lin0768 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH1528
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1528
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-309 <GLA>
A:Cross-references: UNIPROT:Q92DP5; GB:AL592022; PIDN:CAC96000.1; PID:g16413219; GSPDB:B
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0768

Query Match 4.5%; Score 130.5; DB 2; Length 309;
Best Local Similarity 22.3%; Pred. No. 0.0044;
Matches 42; Conservative 33; Mismatches 88; Indels 25; Gaps 5;

QY 134 LVFINPFGKGQKRIYERKVAPLFTLASITTDIIIVTEHANOAKETLYEINIDKVDGIVC 193
Db 5 LLIVNPSSGKEKGK-TYQGTKEVLKKRYDEVRLEKAGDATEFASWASEQGFDAVIA 63

QY 194 VGGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPVPSLSRIGIIPAGSDCVCYSTVGTSD 253
Db 64 MGGDGLNETINGL-----AIHEKRP-----DFGFIPLGTVDNLARSVGIPLK 106


```
Db      162 YKGLEK-----LPQI-----SPTDVR----- 178
Qy      372 LEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLGSPAHLGDGSSDLILIRKCSRFNF 431
Db      179 IEYDGKLFEGEIM--MFLVSN-TNSVGGFER---LAPNASLRDGMDFDIIVKKTSPPEF 231
Qy      432 LRFL-----IRHTNQDQDFDTFVEYRVK 456
Db      232 LHLAALRGEHGHKPK-----VLYVQANRIK 258
```

Search completed: September 3, 2005, 04:38:44
Job time : 25.7411 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 03:56:01 ; Search time 29.3937 Seconds
(without alignments)
1363.781 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 MGATGAEPLOSVLWVKQOR.....QLVRLFARGIENPKPDHS 537

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	801.5	27.8	687	4	US-09-270-767-45874
2	585.5	20.3	359	4	US-09-270-767-46720
3	306	10.6	490	4	US-10-053-510-19
4	302.5	10.5	299	4	US-09-270-767-61405
5	301	10.4	524	4	US-10-053-510-20
6	271.5	9.4	618	4	US-09-970-516-4
7	271.5	9.4	618	4	US-09-817-676A-14
8	267.5	9.3	617	4	US-09-817-676A-12
9	266.5	9.2	384	4	US-09-959-897-2
10	264.5	9.2	384	4	US-09-970-516-2
11	263.5	9.1	384	4	US-09-949-016-7026
12	263.5	9.1	384	4	US-09-796-487-3
13	260	9.0	388	4	US-09-817-676A-15
14	260	9.0	388	4	US-09-796-487-2
15	255.5	8.8	368	4	US-10-053-510-21
16	254	8.8	373	4	US-09-796-487-5
17	254	8.8	381	4	US-09-796-487-1
18	254	8.8	381	4	US-09-796-487-4
19	247.5	8.6	382	4	US-09-970-516-6
20	244.5	8.5	392	4	US-09-796-487-6
21	242	8.4	424	4	US-09-796-487-8
22	228	7.9	536	4	US-09-248-796A-15859
23	206.5	7.2	204	4	US-09-796-487-9
24	198	6.9	403	4	US-09-796-487-7
25	161.5	5.6	312	4	US-09-949-016-9811
26	153	5.3	313	4	US-09-107-532A-7154
27	148	5.1	119	4	US-09-205-258-788

RESULT 1
US-09-270-767-45874
; Sequence 45874, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 45874
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-45874

ALIGNMENTS

28	146	5.1	372	4	US-09-603-208A-264	Sequence 264, App
29	144	5.0	319	4	US-09-270-767-38102	Sequence 38102, A
30	144	5.0	319	4	US-09-270-767-53319	Sequence 53319, A
31	126.5	4.4	293	4	US-09-205-258-328	Sequence 328, App
32	123.5	4.3	304	4	US-09-134-000C-5366	Sequence 5366, App
33	119.5	4.1	283	4	US-09-634-238-373	Sequence 373, App
34	118	4.1	303	4	US-09-107-532A-5133	Sequence 5133, App
35	117	4.1	160	4	US-09-270-767-62339	Sequence 62339, A
36	116.5	4.0	294	4	US-09-583-110-3740	Sequence 3740, App
37	116.5	4.0	315	4	US-09-107-433-3121	Sequence 3121, App
38	113.5	3.9	345	4	US-09-134-000C-3813	Sequence 3813, App
39	110	3.8	324	3	US-09-134-001C-4774	Sequence 4774, App
40	107.5	3.7	325	4	US-09-107-532A-6786	Sequence 6786, App
41	107	3.7	63	4	US-09-959-897-45	Sequence 45, Appl
42	104.5	3.6	64	4	US-09-959-897-37	Sequence 37, Appl
43	103.5	3.6	316	4	US-09-710-279-1350	Sequence 1350, App
44	100	3.5	974	4	US-09-883-134-7	Sequence 7, Appl
45	98.5	3.4	297	4	US-09-134-000C-3616	Sequence 3616, App

Query Match 27.8%; Score 801.5; DB 4; Length 687;
Best Local Similarity 31.3%; Pred. No. 1.6e-74;
Matches 197; Conservative 97; Mismatches 206; Indels 129; Gaps 16;

Qy	11	QSVLWVKQORCAVSL	--PARALLRWRS	PGPGAGACADACSV	PSVIIAVEETDVHGK	68
Db	59	QQLVWERLQIKQSPQNEAKPL	-----PPDSPAQPGGICSGPQSHVLHD	--DVVSI	111	
Qy	69	HQSGKWKQMEKP	-----YAFVHCVKRARR	-----HRKWAQ	101	
Db	112	RSQDTKASSLKPPSPGSRSSGCGSDVAQ	QKPTSQVLTINYAMRLSKSQDTCNRWELRR	171		
Qy	102	VTFWCPBEOLCHLWLTIREMLEKLT	---SRPKLLVFINPFGGKGQKRYERKVALPLF	158		
Db	172	LTFFNSDPYIVRQWDQLQIRLHSSSPTRMKRVRL	LLVFINPYGKKGAGQYERHVRIF	231		
Qy	159	TLASITTDIIIVTEHANOAKETLYEINIDKVD	IGVGVGDGMFSVLHGLIGRTORSAGVD	218		
Db	232	QLAGVDATCITTORANQVKOILLSHDLGVYDA	VCCVGDGTVAEVINGLIIFROMRELGLD	291		
Qy	219	QNHPRVLVPSLRIGIIPAGSTDCVYSTVGT	SDAETSAIHVVGDSLAMDVSSVHHNS	278		
Db	292	EQRPPYIPRP-ALPVGVI	PAGSTDTIAYSMHGTAADVTAATHVLGQHRGLD	350		
Qy	279	TLLRYSVLLGYGYGDI	IKDSEKKRWLGARYDFSLGKTLFSLHHCYEGTVSFL	332		
Db	351	SLLRFCASVLSYGLGDVAAQSEN	YRWMPRRYYSVGKAPLNNRGGDAELRMLEEDLL	410		
Qy	333	---PAQTVGSPRD	-----RKPCRAGCFVC	-----RQSKQOLEEBQKALYGLE	373	
Db	411	LTTPLEDIPQSDVCSLGSVP	SVPCYANCQRCFASSIQBQSRSLFIOESKEA	----E	466	

```
Qy 374 AAEDVE-----EWOVCGKFLAINAT 394
Db 467 RNQVETEDSHLAASEAALLPRPNCNLRPTGSISSMRNLGNDQWKVYVGNFMCIGA 526
Qy 395 NMSCACRRSPGLSPAHLGSSDLILIRKSRFNFLIRHTNQO-QDFDFTFVEVY 453
Db 527 NITCARSPNGISRYSHLGDCLDLILVKKTSLNNVFLNTAGRSGDIRNLPFVEVY 586
Qy 454 RVKKFQFTSKHMEDESDLKEGKKRFQGHIC-----SSHPSCCCTVSSSNWNCDE 504
Db 587 RTRERFRFTFSAEEDYSLAGS-----CQITPPPEMTAHS--STEFSSWNCDE 635
Qy 505 VLHSPAIEVRVHCQLVRIFARGIENPKP 533
Db 636 VVTDLDIRSHCQLIEVFMRGPHYSYKP 664

RESULT 2
US-09-270-767-46720
; Sequence 46720, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 46720
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46720

Query Match 20.3%; Score 585.5; DB 4; Length 359;
Best Local Similarity 35.5%; Pred. No. 2.4e-52;
Matches 125; Conservative 55; Mismatches 103; Indels 69; Gaps 6;

Qy 129 RPKHLLVINPFGGQGGKRIYERKVAFLPLASITTTDIIIVTEHANQAKETLYEINDKY 188
Db 13 RVRLELLVINPYGGRKAGAQTYERHVRPIFQLAGVDATCITTQRANQVKDILLSHDLGVY 72
Qy 189 DGIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRVLVPSLSRIGIIPAGSTDCVCYST 248
Db 73 DAVCCVGSGDGTVAEVINGLIIFQWRELGLDQRPPYIPRP-ALPGVIPAGSTDTIAYSM 131
Qy 249 VGTSDAETSAHLIVVGDSLAMDVSSVHNSLTLLRYSVSLGYPGDIIDKSEKKRWLGL 308
Db 132 HGTADVRTAAIHVILGQHRGLDVCVSNQSLRFCASVLSYGYLGDVAAQSENYRWMP 191
Qy 309 ARYDPSGLKTLFHHCHYEGTVSFL-----PAQHTVGSPPD-----RKPCCAGC 351
Db 192 RRYEYSGVKAFNNRNGYDAELRMLBEPDLLTTPLEDIPQSPDSVCSLGESVPSVCYANC 251
Qy 352 FVC-----RQSKQLEEEOKKALYGLEAAEDVE-----ERNOQVETEDSHLAASEAALLPRPNCNLR 379
Db 252 QRCSPASSIQORSLSLFTQESKEA-----ERNOQVETEDSHLAASEAALLPRPNCNLR 307
Qy 380 -----EWOVCGKFLAINATNMSCACRRSPRLGSPAHLGPG 416
Db 308 LPTGSISSMRNLGNDQWKVYVGNFPMICGANITCACARSPNGISRYSHLGDG 359

RESULT 3
US-10-053-510-19
; Sequence 19, Application US/10053510
; Patent No. 6830881
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; APPLICANT: Fyret, Henrik
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; POLYNUCLEOTIDES AND MODULATING AGENTS AND
```

```
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053,510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-053-510-19

Query Match 10.6%; Score 306; DB 4; Length 490;
Best Local Similarity 24.6%; Pred. No. 6.8e-23;
Matches 119; Conservative 68; Mismatches 165; Indels 132; Gaps 19;

Qy 110 QLCHLWLOTLREMLE-KLTSRP----KHLLVFNPFPGKGQGGKRIYERKVAFLPLASIT 164
Db 12 QTAEMWHHTIRKHKRNGSGSPADCGKQLILLNPKSGSGKRGELFQKQVAPLLTEAEVQ 71
Qy 165 TDIIVTEHANQAKE---TLYEINIDKYDGIVCVGGDMFSEVLHGLIGRTORSAGVDONH 221
Db 72 YDLOITTHPOVAKFVETRRDL-LTRYSGIVVASGGDLFYEVNLGLMER-----MDW 122
Qy 222 PRAVLVPSLSRIGIIPAGSTDCV-----CYSTVGTSDAETSAHLIVVGDSLAMDYSSV 274
Db 123 RRAC---RELPLGIIPCGSGNGLAKSVAAHHCNEPEYEPKPIHLHATLTCMAGKSTPMDVVRV 179
Qy 275 H---HNSTLLRYSVSLGYPGDIIDKSEKKRWLGLARYDFGLKTLFHHCHYEGTVSFP 331
Db 180 ELATRDKHFVMSFLSVGWLIIADIDIESRLRSIGAQRFTLWAIKRLIGLSYKGRVSY 239
Qy 332 -----LPAOHTVGSPPDRKPCRAGCFVCRQSKQLEEEOKKALYGLEAA 375
Db 240 LLGKKEKPEPVEAARELPAESTAGIRSSLPLNAGEP-----HDLPEEEGEAVLDGQFA 295
Qy 376 EDV-----EEQVVC-----GKFLAINATNMSCACR-----RSP 404
Db 296 DAISLDRSVYRQHADSWHSAHSRRRTAYVSLGGPSMRNSRMSISQRIEANAABFAERVP 355
Qy 405 RG-----LSPAHLGDSDDLILIRK- 425
Db 356 TGTIPPLQMLPSLSDGWICEDGDFVMVHAAAYTHLSSDVPFAPESRLDDGLYLVIRRG 415
Qy 426 CSRENEFLRLIRHTNQDQF---DFTFVYVYVKKFOFTSKHMEDESD---LKEGKKR 479
Db 416 VSRHQLNFMFL-NLNAGTHLPIDGEDPFIKVPVPCRAFR-----IEPSSSDGILVVDGERVE 469
Qy 480 FQHI 483
Db 470 YGPI 473

RESULT 4
US-09-270-767-61405
; Sequence 61405, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 61405
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-61405

Query Match 10.5%; Score 302.5; DB 4; Length 299;
Best Local Similarity 29.2%; Pred. No. 7.3e-23;
```



```
Matches 85; Conservative 35; Mismatches 82; Indels 89; Gaps 9;
QY 317 KTFLLSHHCYGVTVSL-----PAQTVGSPRD-----RKPCRACGFCV----- 354
Db 1 KAFLLNRGYDABRLMLEEPDLLTTPLEDIPQSPDSVCSGLBSVPSVCVANCQRCSPASS 60
QY 355 ---ROSKOOLEEOKKALYGLEAAEDVE----- 379
Db 61 IOEQRSSLPFOESKEA-----ERNQOVETEDSHLAASEAALLRPRPRGNLRLPTGSISS 116
QY 380 -----EQVVCGGFLAINATMNSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNF 432
Db 117 MENLNDQWKVVRGNFFMCGANITACARSPNGISRYSHLGDGCLDLILYKKTSLNNV 176
QY 433 RPLIRHTNQ-QDOFTFVYVYRKKFOFTSKHMEDESDSKKEGKKRFGHIC----- 484
Db 177 RPLLTAGRSGDIRLPPFVYVYTRFRFRTPFSASEEDYSLAGS-----CQIPTPPE 228
QY 485 ---SSHPSCCTVSNSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKP 533
Db 229 EMTAHSS---STEFFSWNCDEGEVTVDDLOITMRSHCQLIEVFMRGPHYSYK 276

RESULT 5
US-10-053-510-20
; Sequence 20, Application US/10053510
; Patent No. 6830881
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; APPLICANT: Fyret, Henrik
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053,510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-053-510-20

Query Match 10.4%; Score 301; DB 4; Length 524;
Best Local Similarity 25.2%; Pred. No. 2.5e-22;
Matches 121; Conservative 66; Mismatches 151; Indels 142; Gaps 20;
QY 115 WLQTLREMLEK-----LTSRPKHLVLPINPFQGGKQKRIYERKVAPLFTLASIT 164
Db 17 WYRSURWQLHRTLEEIFVAPTVDERRRVLVLLNPKSGSGDAREVFNNMHVTPVLNEARVP 76
QY 165 TDIIVTEHANOAKETLYEINIDKYGIVCVGGDMFSEVLHGLIGRTQRSAGVDQNHPR 224
Db 77 YDLYTKSNFAIEFLSTRCLDAMCCVAVGGDGLFHEIVNGLLQKQ-----DWAH--- 127
QY 225 VLVPSSLRIGIIPAGSTDCV-----CYS--TVGTSDAETSALHIVVGDSLAMDVSSVHH 276
Db 128 --VLPHLALGIIPCGSGNGLARSIAHCYNKPVLG-----AALTVISGRSPMDVVRQL 179
QY 277 NSTLRYSVLLGYGYFGDIIKDSEKKRWGLARYDFGLKFTFLSHHCYGVTVSPLPAQH 336
Db 180 QSRSL-YSLFSLSGWLISDVDTESIRMLGYQRFTVMTLVLNLTNGRISYLLTDH 238
QY 337 TVGSPRD-----RKPCRACF-----VCROSKOOLREE----- 364
Db 239 EVSSTHSATGYAAQRMRQSSRCNTHIDMLNGPAPIYHSSAEYLPQEFADVISLETSLNQ 298
QY 365 -----QKALY-----GLEAAE-----DV 378
Db 299 SFRSRCDMSLGGSRRSFYYSISESIYHSLADESEFAGLAASLENROQNYGPASELPDL 358
QY 379 EE-----WQVCGKFLAINA---TNMSCACRRSPRGLSPAHLGDGSSDLILIRK-CS 427
```

```
Db 359 NEPLUSEDOGWLVEGEFVMMHAVYQTHLGDICH-----FAPKAQNDGTIYLLIRAGIS 413
QY 428 RFNFLRFLIR---HTNQDQDFTFVEYVYRKKFQFTSKHMEDESD---LKEGKKRFG 481
Db 414 RPHLLSFLYNSSGTHLPESH-DHVKVLVPVRAFR-----LEPYDNHGIIITVDGERVEFG 467

RESULT 6
US-09-970-516-4
; Sequence 4, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-516-4

Query Match 9.4%; Score 271.5; DB 4; Length 618;
Best Local Similarity 30.0%; Pred. No. 4e-19;
Matches 78; Conservative 44; Mismatches 111; Indels 27; Gaps 4;
QY 91 RARHRWKAAQVTFWCPEEQQLCHLWLOTLRMLBKLTSRPHLLVFINPFGGKGGRKIY 150
Db 111 RAEQRWATALTCL-----LRGLPLGDSGITPDLLPRPRLLLLVNPFQGRGLAWQWC 164
QY 151 ERKVAPLFTLASITTDIIVTEHANOAKETLYEINIDKYGIVCVGGDMFSEVLHGLIGR 210
Db 165 KHVLPMLISEAGLSFNLIQTERQNHARELVQGLSLSEWDGIVTVSGDGLLHEVLNGLLDR 224
QY 211 TQRSAGVDQNHPRAVLPSSLRIGIIPAGSTDCVCS-----TVGTSDAETSALHI 261
Db 225 -----PDWEAEVKNP-----VGLPCGSGNALAGAVNQHGFEPALEGLDLLNCSLL 272
QY 262 VVGDSLAMDVSSVHHNSTLLRYSVLLGYGYFGDIIKDSEKKRWGLARYDFSGLKTFPLS 321
Db 273 CRGGHPLDLLSVTLASGRCFSLSVAWGFSVDVIQSERFALGSARFTLGTVLGLAT 332
QY 322 HHCYEGTVSFLPAQHTVGSP 341
Db 333 LHTRYGRLSYLPATVPEPASP 352

RESULT 7
US-09-817-676A-14
; Sequence 14, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-817-676A-14
```

Query Match	9.4%;	Score	271.5;	DB	4;	Length	618;
Best Local Similarity	30.0%;	Pred. No.	4e-19;				
Matches	78;	Conservative	44;	Mismatches	111;	Indels	27;
Gaps	4;						
Qy	91	RARRHRKWKAAQVFWCP	EEQCHLWQLT	RLRMLKLT	SRPKHLVLF	INPFGKGQ	GKRIY 150
Db	111	RABAAQRWATA	LTCL-----	LRGLPL	PCDGEIT	PDLLPR	PRPRLLLVNPF
							GGRGGLAWQC 164
Qy	151	ERKVAPLFT	LTASITTD	IVTEHANQ	AKETLYEIN	DKVDGIV	CGDGMFSEVLH
							IGR 210
Db	165	KNHVLPM	ISEAGLSFN	LIIQTERQ	HNARELV	VGLSLS	EDGWGIVTVSG
							DGLLHEVLNGL
							LLDR 224
Qy	211	TQRSAGVDQ	NHPRAVL	VPSSLRIGI	IIPAGS	PTDCVCS	-----TVGTS
							DAETSALHI 261
Db	225	-----	PDWEAA	VKMP----	VGIL	PCSGNAL	AGAVNQHG
							GFEPALGLDLL
							LLNCSLL 272
Qy	262	VVGDSLAMD	VSSVHHN	TLLRYS	VSLLG	YGFYDII	KDSEKKWGL
							ARYDFSG
							KTFLS 321
Db	273	CRGGGH	PLDLLSV	TLASG	RCFCF	SLFVAM	GFVSVDIQ
							QSERFRAL
							GSARFTL
							GTVLGIAT 332
Qy	322	HHCYEGT	VSFELP	QAHTV	GSP		341
Db	333	LHTYGR	LSYLPAT	VEPASP			352

```

RESULT 8
US-09-817-676A-12
; Sequence 12, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; TITLE OF INVENTION: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 12
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-817-676A-12

```

```

RESULT 9
US-09-959-897-2
; Sequence 2, Application US/09959897
; Patent No. 6730480
; GENERAL INFORMATION:
; APPLICANT: PITSON, Stuart M
; APPLICANT: Brian, WATTENBERG W
; APPLICANT: Pu, XIA
; APPLICANT: Richard, D'ANDREA J
; APPLICANT: Jennifer, BAMBLE R
; APPLICANT: Mathew, VADAS A
; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
; FILE REFERENCE: PITSON=1
; CURRENT APPLICATION NUMBER: US/09/959,897
; PRIORITY FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00457
; PRIORITY FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: AU PQ 0339
; PRIORITY FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: AU PQ 1504
; PRIORITY FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-959-897-2

```

RESULT 10
US-09-970-516-2
; Sequence 2, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of b

; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases

```
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-516-2

Query Match
Best Local Similarity 9.2%; Score 264.5; DB 4; Length 384;
Matches 107; Conservative 68; Mismatches 169; Indels 79; Gaps 16;

QY 129 RPKHLVFINPFGGKGKRIYERKVAFLFTLASITTDIIIVTEHANQAKETLYEINIDKY 188
Db 13 RPCRVLNLPNGGKGKALQFRSHVQPLLAEEISFTMLTERNNHARELVRSELGRW 72

QY 189 DGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVPSLSRIGIIPAGSTDCVC--- 245
Db 73 DALVMSGDGLMHEVVGMLMERPDWETAIQK-----PLCSLPAGSGNALAASL 120

QY 246 -----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVLSLLGYGYGDIIDK 299
Db 121 NHYAGYEQVTNEDLLTNTCLLCRRLLSPMNLISLTASGLRFSVLGLAWGFIADVDLE 180

QY 300 SEKKRWGLARYDFSLGKTFLSHCYEGTVSFLPAQHTVGSPRDKPCRAGFCVCRQSKQ 359
Db 181 SEKYRRLGEMRTLTGTLRLAALRYRGLAYLPVGR-VGSKTPASP-----VVVQGPV 234

QY 360 QLEBEQKALYGLEAAEDV-BEWOVVCVK-----FLAINATNMSCACRRSPRGLSPAHLG 414
Db 235 DAH-----LVPLE--EPVPSHTVVPDEDFVLVLLALLSHLGSEMFAPMGRCAA--- 282

QY 415 DGSSDLILIRK-CSRFNPLRLI-----RHTNQDQDFTFVEVYRVKKFQFTSKHMEDE 472
Db 283 -GVMEHLFYVRAGVSRAMLRLFLAMEKGRHM-----EVECPYLIVYVVAFR-----LEPKD--- 333

QY 473 KEGGKKRFHGHCSSHPSCCTVSNSSWNCDBVLHSPAIEVRVHCOLVRLPARGIEENP- 531
Db 334 ---GKGVFA-----VDGELMVSEAVQGVHPNPFYFMVSGCGVEPPPS 371

QY 532 -KP 533
Db 372 WKP 374
```

RESULT 11

```
US-09-949-016-7026
; Sequence 7026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7026
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7026
```

```
Query Match
Best Local Similarity 9.1%; Score 263.5; DB 4; Length 384;
Matches 107; Conservative 69; Mismatches 164; Indels 87; Gaps 17;

QY 129 RPKHLVFINPFGGKGKRIYERKVAFLFTLASITTDIIIVTEHANQAKETLYEINIDKY 188
Db 13 RPCRVLNLPNGGKGKALQFRSHVQPLLAEEISFTMLTERNNHARELVRSELGRW 72

QY 189 DGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVPSLSRIGIIPAGSTDCVC--- 245
Db 73 DALVMSGDGLMHEVVGMLMERPDWETAIQK-----PLCSLPAGSGNALAASL 120

QY 246 -----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVLSLLGYGYGDIIDK 299
Db 121 NHYAGYEQVTNEDLLTNTCLLCRRLLSPMNLISLTASGLRFSVLGLAWGFIADVDLE 180

QY 300 SEKKRWGLARYDFSLGKTFLSHCYEGTVSFLPAQHTVGSPRDKPCRAGFCVCRQSKQ 359
Db 181 SEKYRRLGEMRTLTGTLRLAALRYRGLAYLPVGR-VGSKTPASP-----VVVQGPV 234

QY 360 QLEBEQKALYGLEAAEDV-BEWOVVCVK-----FLAINATNMSCACRRSPRGLSPAHLG 414
Db 235 DAH-----LVPLE--EPVPSHTVVPDEDFVLVLLALLSHLGSEMFAPMGRCAA--- 282

QY 415 DGSSDLILIRK-CSRFNPLRLI-----RHTNQDQDFTFVEVYRVKKFQFTSKHMEDE 468
Db 283 -GVMEHLFYVRAGVSRAMLRLFLAMEKGRHM-----EVECPYLIVYVVAFR-----LEPK 332

QY 469 DSDLKEGKKGFGHCSSHPSCCTVSNSSWNCDBVLHSPAIEVRVHCOLVRLPARGIE 528
Db 333 D-----GKGVFA-----VDGELMVSEAVQGVHPNPFYFMVSGCGVE 367

QY 529 ENP--KP 533
Db 368 PPSWKP 374
```

RESULT 12

```
US-09-796-487-3
; Sequence 3, Application US/09796487
; Patent No. 6830916
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(384)
; OTHER INFORMATION: seq id NO 3 is the peptide sequence of hSPHK1 in Fig. 3, correspo
; nding to amino acid residue 1 to 384 of Homo sapiens SPHK-1 of Ge
; nBank sequence Accession Number AAF73423.
; OTHER INFORMATION: nBank sequence Accession Number AAF73423.
; PUBLICATION INFORMATION:
; AUTHORS: Nava et al.
; TITLE: Functional characterization of human sphingosine kinase-1
; JOURNAL: FEBS Lett.
; VOLUME: 473
; ISSUE: 1
; PAGES: 81-84
; DATE: 2000
; DATABASE ACCESSION NUMBER: AAF73423
```



```
ISSUE: 37
PAGES: 23722-23728
DATE: 1998
DATABASE ACCESSION NUMBER: AAC61698
DATABASE ENTRY DATE: 1998-09-26
RELEVANT RESIDUES: (1)..(388)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AAC61698
DATABASE ENTRY DATE: 1998-09-26
RELEVANT RESIDUES: (1)..(388)
US-09-796-487-2

Query Match
Best Local Similarity 23.3%; Pred. No. 3.1e-18; Length 388;
Matches 100; Conservative 70; Mismatches 17; Indels 86; Gaps 13;

QY 104 FWCPEQCHLWQLTREMLEKLTSPKHLIVFINFGKGQKRIYERKVAFLFTLASI 163
Db 2 WWC-----CVLFVV---ECPRGLLPFCRVLVLPNPGKGKALQLFQSRVQPFLEAEI 53

QY 164 TTDIIIVTEHANOAKETLYEINIDKYDGIYVCGDGMFSEVLHGLIGRTQRSAGVDQNHPR 223
Db 54 TFKLILTERKHNARELVCAEELGHWDALAVMSGDGLMHEVVNGLMERPDWETAIOK---- 109

QY 224 AVLVPSSLRIGIIPAGSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDYSSV 274
Db 110 -----PLCSLPGGSGNALAASVNHVAGYEQVTNEDLLINCTLLCCRRLLSPMNLISL 161

QY 275 HNNSTLLRYSLLYGYFYGDIIKDSEKKRWGLARYDFGLKTFSLHHCYEGTVSFLPA 334
Db 162 HTASGLRLYSLVLSWGFVADVLESEKYRLGEIRFVGTFFRLASLRIVQGLAYLP- 220

QY 335 QHTVGSPRDRKPCRAGCFVCRSQKQLEEBEOKKALYGLEAAEDV-BEQVVVCGK----FL 389
Db 221 ---VGTASKRPA-----STLVQKGPVDTHLVPLEEPVPSHWTVPVEQDFVLVL 266

QY 390 AINATNMSCACRRSPRGLSPAHLGDGSSDLILIRK-CSRFNFLR-FLIRHTNQDQDF 447
Db 267 VLLHTLHLSSELFAAPMGRCFA-----GYMHLFYVRAGVSRALLRLFLAMQKGMELDC 321

QY 448 TFVEYRVYKFTQTSKHMEDESDLEKGGKKRFGHICSSHPSCCCTVSNSSWNCDCGEVLH 507
Db 322 PYLHVVPVAFRLBPRS-----QGVFVSDGELMV 351

QY 508 SPAIEVRVH 516
Db 352 CEAVQGVH 360
```

```
RESULT 15
US-10-053-510-21
Sequence 21, Application US/10053510
Patent No. 6830881
GENERAL INFORMATION:
APPLICANT: Saba, Julie D.
APPLICANT: Fyrest, Henrik
TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
TITLE OF INVENTION: METHODS OF USE THEREFOR
FILE REFERENCE: 200116.402C2
CURRENT APPLICATION NUMBER: US/10/053.510
CURRENT FILING DATE: 2002-01-17
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 368
TYPE: PRT
ORGANISM: Homo sapiens
US-10-053-510-21

Query Match
Best Local Similarity 8.8%; Score 255.5; DB 4; Length 368;
Matches 105; Conservative 68; Mismatches 167; Indels 79; Gaps 16;
```

```
QY 133 LLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKYDGI 192
Db 1 VLVLLNPRGGKALQLFRSHVQPLLAETISFTMLTERRNHARELVRSBELGRWDALV 60

QY 193 CVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSPSLRIGIIPAGSTDCVC----- 245
Db 61 VMSGDGLMHEVVNGLMERPDWETAIOK-----PLCSLPGGSGNALAASLNHYA 108

QY 246 -YSTVGTSDAETSALHIVVGDLSA-MDYSSVHNNSTLLRYSVLSLIGYFYGDIIKDSEKK 303
Db 109 GYEQVTNEDLLTNTCTLLCCRRLLSPMNLISLHTASGLRLFSVLSLANGFIADVLESEKY 168

QY 304 RWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRSQKQLEE 363
Db 169 RRLGEMFTLTGTFURLAALRTYRGLAYLPVGR-VGSKTPASP-----VVVQGPVDAH- 221

QY 364 EQKKALYGLEAAEDV-BEQVVVCGK----FLAINATNMSCACRRSPRGLSPAHLGDGSS 418
Db 222 -----LVPLE--EPVPSHWTVPDEDFVLVLLALLHSLGSEMFAAPMGRCFA-----GVM 269

QY 419 DLILIRK-CSRFNFLR-FLIRHTNQDQDFTFVEYRVYKFTQTSKHMEDESDLEKGG 476
Db 270 HLFYVRAGVSRALLRLFLAMEKGRHMEYECPLYLVYVVAFR-----LEPKD-----G 318

QY 477 KKRFGHICSSHPSCCCTVSNSSWNCDCGEVLHSPAIEVRVHCOLVRLFARGIEENP--KP 533
Db 319 KGVEA-----VDGELMVSEAVQGVHFNYPFMWVSGCVPEPPPSWKP 358
```

Search completed: September 3, 2005, 04:10:36
Job time : 30.3937 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:09:18 ; Search time 109.661 Seconds
(without alignments)
1893.930 Million cell updates/sec

Title: US-10-631-958-10
Perfect score: 2888
Sequence: 1 MGATCAAPLQSLVWVKQR.....QLVRLFARGIENKPDPSHS 537

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A: Geneseq 16Dec04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2888	100.0	537	ABB07856	Abb07856 Human sph
2	2888	100.0	537	ADA05880	Ada05880 Human NOV
3	2888	100.0	537	ADJ96664	Adj96664 Human lip
4	2888	100.0	537	ADN62845	Adn62845 Human NOV
5	2888	100.0	562	ABB07857	Abb07857 Human sph
6	2880	99.7	537	AAM49115	Aam49115 Human Cer
7	2588	89.6	481	ABR56302	AbR56302 Human Sph
8	2463	85.3	460	AAY96059	Aay96059 Human sph
9	2456.5	85.1	471	AAE07884	Aae07884 Human sph
10	2428	84.1	531	ADP55248	Adp55248 Human PRO
11	2350	81.4	536	ADS11054	Ads11054 Human the
12	2210	76.5	416	ABR56301	AbR56301 Human Sph
13	2208.5	76.5	746	ABG13541	Abg13541 Novel hum
14	1714.5	59.4	727	ABG13543	Abg13543 Novel hum
15	1640.5	56.8	326	ABB07854	Abb07854 Human sph
16	1361.5	47.1	454	ABG13544	Abg13544 Novel hum
17	1055	36.5	228	ABR41822	AbR41822 Human ORF
18	1035	35.8	255	ADS12267	Ads12267 Human the
19	1032	35.7	190	ABP64313	Abp64313 Human pro
20	986	34.1	182	ABR42383	AbR42383 Human ORF
21	675	23.4	596	ABB69669	Abb69669 Drosophil
22	664	23.0	136	ABG13540	Abg13540 Novel hum
23	647	22.4	144	AAE07885	Aae07885 Partial r
24	628	21.7	136	ABG13542	Abg13542 Novel hum
25	582.5	20.2	532	ADQ88891	Adq88891 Novel hum

26	335	11.6	79	4	AAE07886	Aae07886 Partial m
27	310.5	10.8	641	4	ABB58465	Abb58465 Drosophil
28	310.5	10.8	641	6	ABR82392	AbR82392 D. melano
29	306	10.6	490	6	ABR82389	AbR82389 D. melano
30	301	10.4	524	6	ABR82390	AbR82390 D. melano
31	301	10.4	907	4	ABB57980	Abb57980 Drosophil
32	301	10.4	907	6	ABR82393	AbR82393 D. melano
33	277.5	9.6	618	4	AAU09075	Aau09075 Human sph
34	276	9.6	806	4	ABG21144	Abg21144 Novel hum
35	271.5	9.4	618	4	AAU09074	Aau09074 Human sph
36	271.5	9.4	618	5	ABG31587	Abg31587 Human sph
37	271.5	9.4	618	5	ABB07855	Abb07855 Protein s
38	271.5	9.4	618	7	ADE38417	AdE38417 Human pro
39	271.5	9.4	618	8	ADQ15182	AdQ15182 Human can
40	271.5	9.4	654	4	ABU52806	Abu52806 Human sig
41	271.5	9.4	654	8	ADJ66569	Adj66569 Sphingos
42	267.5	9.3	617	4	AAU09073	Aau09073 Mouse sph
43	266.5	9.2	384	4	AAB48007	Aab48007 Human sph
44	266.5	9.2	384	7	ADF28783	Adf28783 Human sph
45	264.5	9.2	384	3	AAAB18659	AaB18659 A human r

ALIGNMENTS

RESULT 1
ABB07856
ID ABB07856 standard; protein; 537 AA.
XX AC ABB07856;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human sphingosine kinase-like protein.
XX
KW Human sphingosine kinase-like protein; intracellular signalling;
cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
autoimmune disease; rheumatoid arthritis; Parkinson's disease.
XX
OS Homo sapiens.
XX
PN WO200228906-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-EP011516.
XX
PR 06-OCT-2000; 2000US-0238005P.
23-AUG-2001; 2001US-0314113P.
(FARB) BAYER AG.
Kossida S, Encinas J;
WPI; 2002-340094/37.
N-PSDB; ABL40828.
New reagent for modulating the activity of sphingosine kinase-like protein polypeptide or polynucleotide and treating cancer, asthma, allergy, an autoimmune disease, or a central or peripheral nervous system disorder.
Claim 25; Fig 10; 120pp; English.

The invention relates to a human sphingosine kinase-like protein. The polypeptide can be expressed by standard recombinant methodology. The sphingosine kinase-like protein and gene can be used to regulate intracellular signalling and consequently cell proliferation and apoptosis. Such regulation is useful for treating cancer, allergies (e.g. asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and peripheral nervous system disorders (e.g. Parkinson's disease). The present sequence represents the human sphingosine kinase-like protein

[illegible]


```
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence represents a human NOVX from the
CC present invention.
XX
SQ Sequence 537 AA;

Query Match      100.0%; Score 2888; DB 6; Length 537;
Best Local Similarity 100.0%; Pred. No. 2e-274;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLOSIVLVKQRCVSLPARALLRWRSPGAGAGADACSPVSEIITAV 60
Db 1 MGATGAAPLOSIVLVKQRCVSLPARALLRWRSPGAGAGADACSPVSEIITAV 60

Qy 61 BETDVHGKHQSGKWKQKPKYAFVTVCKRARRRHWKWAQVTFWCPEQLCHLWLQTLR 120
Db 61 BETDVHGKHQSGKWKQKPKYAFVTVCKRARRRHWKWAQVTFWCPEQLCHLWLQTLR 120

Qy 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIVTEHANOAKETL 180
Db 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIVTEHANOAKETL 180

Qy 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRVLPSSLRIGIIPAGS 240

Qy 241 TDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLGIGFYGDIIKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLGIGFYGDIIKDS 300

Qy 301 EKRWGLGARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQQ 360
Db 301 EKRWGLGARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQQ 360

Qy 361 LEEBQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 361 LEEBQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420

Qy 421 ILIRKCSRFNRLIRHTNQDQDFTFVEYVRVKKFQFTSKHMEDESDLKEGGKKRF 480
Db 421 ILIRKCSRFNRLIRHTNQDQDFTFVEYVRVKKFQFTSKHMEDESDLKEGGKKRF 480

Qy 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537

RESULT 3
ADJ96664
ID ADJ96664 standard; protein; 537 AA.
XX
AC ADJ96664;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human lipid kinase KIAA1646 protein SeqID 121.
XX
KW kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
KW PTK; Stk; gene therapy; cancer; immune-related disease;
KW cardiovascular disease; brain; neuronal associated disease; metabolic;
KW inflammatory disorder; cytostatic; neuroprotective; immunomodulator;
KW antiinflammatory; enzyme; lipid kinase; KIAA1646.
XX
OS Homo sapiens.
OS 44.
XX
PN WO2004006838-A2.
XX
PD 22-JAN-2004.
XX
```

```
PF 15-JUL-2003; 2003WO-US021730.
XX
PR 15-JUL-2002; 2002US-0395632P.
XX
PA (SUGE-) SUGEN INC.
XX
PI Whyte D, Manning G, Caenepeel S;
XX
DR WPI; 2004-122753/12.
XX
DR N-PSDB; ADJ96598.
XX
PT New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX
PS Claim 1; SEQ ID NO 121; 366pp; English.
XX
CC This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytosatic, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polypeptide sequence is a human kinase protein sequence
CC of the invention.
XX
SQ Sequence 537 AA;

Query Match      100.0%; Score 2888; DB 8; Length 537;
Best Local Similarity 100.0%; Pred. No. 2e-274;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLOSIVLVKQRCVSLPARALLRWRSPGAGAGADACSPVSEIITAV 60
Db 1 MGATGAAPLOSIVLVKQRCVSLPARALLRWRSPGAGAGADACSPVSEIITAV 60

Qy 61 BETDVHGKHQSGKWKQKPKYAFVTVCKRARRRHWKWAQVTFWCPEQLCHLWLQTLR 120
Db 61 BETDVHGKHQSGKWKQKPKYAFVTVCKRARRRHWKWAQVTFWCPEQLCHLWLQTLR 120

Qy 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIVTEHANOAKETL 180
Db 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIVTEHANOAKETL 180

Qy 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRVLPSSLRIGIIPAGS 240

Qy 241 TDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLGIGFYGDIIKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLGIGFYGDIIKDS 300

Qy 301 EKRWGLGARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQQ 360
Db 301 EKRWGLGARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQQ 360

Qy 361 LEEBQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 361 LEEBQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420

Qy 421 ILIRKCSRFNRLIRHTNQDQDFTFVEYVRVKKFQFTSKHMEDESDLKEGGKKRF 480
Db 421 ILIRKCSRFNRLIRHTNQDQDFTFVEYVRVKKFQFTSKHMEDESDLKEGGKKRF 480

Qy 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537
```


Db 61 EETDVHGHQSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQLCHLWLTQLR 120
QY 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180
QY 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDONHPRAVLVPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDONHPRAVLVPSSLRIGIIPAGS 240
QY 241 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLILGYGYGDIIDKDS 300
Db 241 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLILGYGYGDIIDKDS 300
QY 301 EKKRWGLIARYDFSLGKTLFSLHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGFCVCRQSKQ 360
Db 301 EKKRWGLIARYDFSLGKTLFSLHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGFCVCRQSKQ 360
QY 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
Db 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
QY 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKKEQFTSKHMEDESDLKEGGKKRF 480
Db 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKKEQFTSKHMEDESDLKEGGKKRF 480
QY 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 537
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 537

RESULT 5
ABB07857
ID ABB07857 standard; protein; 562 AA.
AC ABB07857;
XX
XX
DT 03-JUL-2002 (first entry)
DE Human sphingosine kinase-like protein.
DE Human sphingosine kinase-like protein; intracellular signalling;
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease.
XX
XX Homo sapiens.
XX
XX WO200228906-A2.
XX
XX 11-APR-2002.
XX
XX 05-OCT-2001; 2001WO-EP011516.
XX
XX 06-OCT-2000; 2000US-0238005P.
XX
XX 23-AUG-2001; 2001US-0314113P.
XX
XX (FARB) BAYER AG.
XX
XX Kossida S, Encinas J;
XX
XX PI
XX
XX WPI; 2002-340094/37.
XX
XX N-P5DB; ABL40828.
XX
XX
PT New reagent for modulating the activity of sphingosine kinase-like
PT protein polypeptide or polynucleotide and treating cancer, asthma,
PT allergy, an autoimmune disease, or a central or peripheral nervous system
PT disorder.
XX
XX
XX Claim 25; Fig 11; 120pp; English.
XX
XX
XX The invention relates to a human sphingosine kinase-like protein. The
CC polypeptide can be expressed by standard recombinant methodology. The
CC sphingosine kinase-like protein and gene can be used to regulate

CC intracellular signalling and consequently cell proliferation and
CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
CC peripheral nervous system disorders (e.g. Parkinson's disease). The
CC present sequence represents the human sphingosine kinase-like protein
XX
XX Sequence 562 AA;
QY
Query Match 100.0%; Score 2888; DB 5; Length 562;
Best Local Similarity 100.0%; Pred. No. 2.2e-274;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGATGAAPLQSVLWVKQRCANVLEPARALLRWRSPGCGAGAGDADACVPVSEIITAV 60
Db 26 MGATGAAPLQSVLWVKQRCANVLEPARALLRWRSPGCGAGAGDADACVPVSEIITAV 85
QY 61 EETDVHGHQSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQLCHLWLTQLR 120
Db 86 EETDVHGHQSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQLCHLWLTQLR 145
QY 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180
Db 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 205
QY 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDONHPRAVLVPSSLRIGIIPAGS 240
Db 206 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDONHPRAVLVPSSLRIGIIPAGS 265
QY 241 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLILGYGYGDIIDKDS 300
Db 266 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLILGYGYGDIIDKDS 325
QY 301 EKKRWGLIARYDFSLGKTLFSLHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGFCVCRQSKQ 360
Db 326 EKKRWGLIARYDFSLGKTLFSLHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGFCVCRQSKQ 385
QY 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
Db 386 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 445
QY 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKKEQFTSKHMEDESDLKEGGKKRF 480
Db 446 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKKEQFTSKHMEDESDLKEGGKKRF 505
QY 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 537
Db 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 562

RESULT 6
AAM49115
ID AAM49115 standard; protein; 537 AA.
XX
XX
AC AAM49115;
XX
XX 20-MAY-2002 (first entry)
XX
XX Human ceramide kinase hCERK1.
XX
XX Human; ceramide kinase; hCERK1; drug screening; gene therapy;
KW neurological disease; inflammation; human immunodeficiency virus;
KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;
KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;
KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;
KW enzyme.
XX
XX Homo sapiens.
XX
XX WO200196575-A1.
XX
XX 20-DEC-2001.
XX
XX 11-JUN-2001; 2001WO-JP004889.

```
XX 14-JUN-2000; 2000JP-00178039.
XX (SANY ) SANKYO CO LTD.
XX Sugiura M, Kono K, Kohama T;
XX WPI; 2002-179513/23.
XX N-PSDB; ABA96945.
XX Human ceramide kinase gene and the enzyme encoded by it for screening
XX substances as drugs for neurological, inflammatory and other disorders.
XX Claim 1; Page 54-57; 61pp; Japanese.
XX This sequence represents a human ceramide kinase designated hCERK1. The
XX invention relates to hCERK1, nucleic acids encoding it, expression
XX vectors and host cells containing hCERK1 nucleic acids, the recombinant
XX production of hCERK1 and antibodies specific for hCERK1. The invention
XX also encompasses methods of isolating hCERK1 from samples, the use of
XX hCERK1 in drug screening, and the use of hCERK1 nucleic acid sequences in
XX gene therapy. hCERK1 mediates the ATP-dependent 1-phosphorylation of
XX ceramides and can be used to screen for therapeutic and preventive agents
XX for a wide range of disorders. Such disorders include neurological
XX disease, inflammation, human immunodeficiency virus (HIV) infection, type
XX 2 diabetes, obesity, sepsis, arteriosclerosis and cancer
XX Sequence 537 AA;
Query Match 99.7%; Score 2880; DB 5; Length 537;
Best Local Similarity 99.6%; Pred. No. 1.2e-273;
Matches 535; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MGATGAEPLOSIVLVKQORCAVSLPARALLRWSPGCGAGAPGADACVPVSEIIAV 60
Db 1 MGATGAEPLOSIVLVKQORCAVSLPARALLRWSPGCGAGAPGADACVPVSEIIAV 60
Qy 61 EETDVHGKQSGKQWKQMEKPYAFTVHCVKRARRHRKWAQVTFWCPEEQCHLWLQTLR 120
Db 61 EETDVHGKQSGKQWKQMEKPYAFTVHCVKRARRHRKWAQVTFWCPEEQCHLWLQTLR 120
Qy 121 EMLEKLTSRPKHLLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANQAKETL 180
Db 121 EMLEKLTSRPKHLLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANQAKETL 180
Qy 181 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Qy 241 TDCVCYSTVGTSDAETSAHIIWVGDLSAMDVSSVHNSLTLLRYSVSLGFGYGDIIKDS 300
Db 241 TDCVCYSTVGTSDAETSAHIIWVGDLSAMDVSSVHNSLTLLRYSVSLGFGYGDIIKDS 300
Qy 301 EKKRWGLGARYDFGLKTLFSLHHCYEGTGVFLPAQHTVGSPPDRKPCRACFCVCRQSKQ 360
Db 301 EKKRWGLGARYDFGLKTLFSLHHCYEGTGVFLPAQHTVGSPPDRKPCRACFCVCRQSKQ 360
Qy 361 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 361 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Qy 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVVEVVRVKKFQFTSKHMEDESDSLKEGKKRF 480
Db 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVVEVVRVKKFQFTSKHMEDESDSLKEGKKRF 480
Qy 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSSH 537
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSSH 537
RESULT 7
ABR56302
ID ABR56302 standard; protein; 481 AA.
```

```
XX ABR56302;
XX 20-NOV-2003 (first entry)
XX Human Sphingosine kinase 4-related protein.
XX Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;
XX platelet transfusion; platelet stabiliser.
XX Homo sapiens.
XX WO2003031627-A1.
XX 17-APR-2003.
XX 28-SEP-2001; 2001WO-JP008537.
XX 28-SEP-2001; 2001WO-JP008537.
XX (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.
XX (CHBI-) CHEM BIOLOGY INST.
XX Igarashi Y, Kihara A;
XX WPI; 2003-354917/33.
XX Platelet derived polypeptides with sphingosine kinase activity for
XX treatment of sphingosine related disorders.
XX Example 4; Fig 4; 39pp; Japanese.
XX The present invention relates to human sphingosine kinase 4 (SPHK4;
XX ABR56301). The kinase can be used for the diagnosis and treatment of
XX sphingosine related disorders. The kinase can also be potentially used
XX for controlling toxicity of platelet transfusion and as a platelet
XX stabiliser. The present sequence was used to illustrate the invention
XX Sequence 481 AA;
Query Match 89.6%; Score 2588; DB 6; Length 481;
Best Local Similarity 100.0%; Pred. No. 5.4e-245;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 57 IIAVEETDVHGKQSGKQWKQMEKPYAFTVHCVKRARRHRKWAQVTFWCPEEQCHLWL 116
Db 1 IIAVEETDVHGKQSGKQWKQMEKPYAFTVHCVKRARRHRKWAQVTFWCPEEQCHLWL 60
Qy 117 QTLREMLEKLTSRPKHLLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANQA 176
Db 61 QTLREMLEKLTSRPKHLLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANQA 120
Qy 177 KETLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGI 236
Db 121 KETLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGI 180
Qy 237 PAGSTDCVCYSTVGTSDAETSAHIIWVGDLSAMDVSSVHNSLTLLRYSVSLGFGYGD 296
Db 181 PAGSTDCVCYSTVGTSDAETSAHIIWVGDLSAMDVSSVHNSLTLLRYSVSLGFGYGD 240
Qy 297 IKDSEKKRWGLGARYDFGLKTLFSLHHCYEGTGVFLPAQHTVGSPPDRKPCRACFCVCRQ 356
Db 241 IKDSEKKRWGLGARYDFGLKTLFSLHHCYEGTGVFLPAQHTVGSPPDRKPCRACFCVCRQ 300
Qy 357 SKQOLEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDG 416
Db 301 SKQOLEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDG 360
Qy 417 SSDLLILIRKCSRFNRLFLIRHTNQDQDFTFVVEVVRVKKFQFTSKHMEDESDSLKEGG 476
Db 361 SSDLLILIRKCSRFNRLFLIRHTNQDQDFTFVVEVVRVKKFQFTSKHMEDESDSLKEGG 420
Qy 477 KKRFGHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSSH 536
```

```

Db 421 KRFEGHICSSHPSCCTVSNWNCDEVLHSPAIEVRVHCOLVRLFARGIEENKPDUSH 480
Qy 537 S 537
Db 481 S 481

RESULT 8
AA96059
ID AA96059 standard; protein; 460 AA.
AC AA96059;
XX
DT 05-DEC-2000 (first entry)
DE Human sphingosine kinase C.
KW Sphingosine kinase C; SKC; human; drug screening; infection;
KW antiinflammatory; antiallergic; anticancer; inflammation; allergy;
KW cancer; therapy; diagnosis.
XX
OS Homo sapiens.
XX
PN WO200052173-A2.
XX
PD 08-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-CA000223.
XX
PR 02-MAR-1999; 99US-0122516P.
XX
PA (ALLX ) NPS ALLELIX CORP.
XX
PI Munroe D, Gupta A, Falzone GR;
XX
DR WPI; 2000-572185/53.
XX
DR N-PSDB; AAA50510.
XX
New human sphingosine kinase A, B and C polynucleotides and polypeptides
useful in e.g. chromosome and gene mapping, and detecting inflammation or
disease associated with abnormal levels of sphingosine kinase expression.
XX
PS Disclosure; Fig 9; 8lpp; English.
XX
XX
The present sequence is that of human sphingosine kinase C (SKC), an
enzyme that phosphorylates sphingosine to form sphingosine 1-phosphate.
The sequence was deduced from that of a polynucleotide (see AAA50510)
isolated from an HeLa cDNA library. The invention provides
polynucleotides (see AAA50508-10) and polypeptides (see AA96057-59) for
the human sphingosine kinase (SK) homologues SKA, SKB and SKC. The
polypeptides can be obtained using recombinant DNA methods, and host
cells containing expression vectors including SK polynucleotides are used
in a claimed method of screening for compounds that inhibit or activate
human SK activity. Human SK specific antibodies, inhibitors, ligands or
their analogues can be used as bioactive agents to treat inflammation or
disease including viral, bacterial or fungal infections, allergic
responses, mechanical injury associated with trauma, hereditary diseases,
lymphoma or carcinoma, and other conditions with activate the genes of
kidney, lung, heart, lymphoid or tissues of the nervous system
XX
SQ Sequence 460 AA;
Query Match 85.3%; Score 2463; DB 3; Length 460;
Best Local Similarity 99.6%; Pred No. 9, 9e-233;
Matches 458; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 78 MEKPYAFTVHCVKARRHRKWAQVTFWCPEQLCHLWLQTLREMLEKLTSPKHLVFI 137
Db 1 MEKPYAFTVHCVKARRHRKWAQVTFWCPEQLCHLWLQTLREMLEKLTSPKHLVFI 60
Qy 138 NPFSGKGQKRIYERKVAFLTLASITTDIIVTEHANOAKETLYEINIDKDYGVCGD 197

```

```

Db 61 NPFSGKGQKRIYERKVAFLTLASITTDIIVTEHANOAKETLYEINIDKDYGVCGD 120
Qy 198 GMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDVCYCVSTGTSDAETS 257
Db 121 GMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDVCYCVSTGTSDAETS 180
Qy 258 ALHIWVGDSLAMDYSSVHHNSTLLRYSVSLIGYFYGDIIKDSKKRWGLGARYDFSLK 317
Db 181 ALHIWVGDSLAMDYSSVHHNSTLLRYSVSLIGYFYGDIIKDSKKRWGLGARYDFSLK 240
Qy 318 TFLSHHCVEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSKQOLEEEOKKALYGLEAAED 377
Db 241 TFLSHHCVEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSKQOLEEEOKKALYGLEAAED 300
Qy 378 VBEQVVCGRKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNLFRLIR 437
Db 301 VBEQVVCGRKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNLFRLIR 360
Qy 438 HTNOODQDFTFVEVYRVKPKQFTSKMEDSDSLKEGKKRFGHICSSHPSCCCTVSN 497
Db 361 HTNOODQDFTFVEVYRVKPKQFTSKMEDSDSLKEGKKRFGHICSSHPSCCCTVSN 420
Qy 498 SWNCDEVLHSPAIEVRVHCOLVRLFARGIEENKPDUSH 537
Db 421 SWNCDEVLHSPAIEVRVHCOLVRLFARGIEENKPDUSH 460

RESULT 9
AAE07884
ID AAE07884 standard; protein; 471 AA.
AC AAE07884;
XX
DT 01-NOV-2001 (first entry)
DE Human sphingosine kinase (SphK) protein #2.
KW Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy;
KW antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
KW cytosatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
KW leukaemia; vasotropic; cell proliferative disorder; vascular disease.
XX
OS Homo sapiens.
XX
PN WO200160990-A2.
XX
PD 23-AUG-2001.
XX
PF 14-FEB-2001; 2001WO-US004789.
XX
PR 14-FEB-2000; 2000US-0182360P.
XX
PR 22-MAR-2000; 2000US-0191261P.
XX
PA (CURA-) CURAGEN CORP.
PA (GETH ) GENENTECH INC.
XX
Rastelli L;
XX
WPI; 2001-514770/56.
XX
N-PSDB; AAD14426.
XX
An isolated Sphingosine kinase polypeptide useful for treating a SphK-
associated disorder especially cancer, restenosis or ischemia in a human.
XX
Claim 1; Fig 1; 107pp; English.
XX
The present invention relates to sphingosine kinase (SphK) polypeptides
and nucleic acids encoding them. SphK is useful for treating a SphK-
associated disorder especially cancers such as leukaemia, lymphoma,
ovarian, breast, lung, colon, testicular, stomach and skin,
atherosclerosis, restenosis or ischaemia and cell proliferative disease
or disorder associated with vascular diseases. SphK gene is used in gene
therapy and antisense-therapy. Sphingolipids serving as signalling

```


QY 361 LEEQKALYGLAEDVEVQVCGKFLAINATMNSCACRSPRGLSPAHLGDGSSDL 420
DB 361 LEEQKALYGLAEDVEVQVCGKFLAINATMNSCACRSPRGLSPAHLGDGSSDL 420
QY 421 ILIRKCSRNFRLIRHTNQDQDFTFVEVYRVKKFOFTSKHMEDESDLKEGGKRF 480
DB 421 ILIRKCSRNFRLIRHTNQDQDFTFVEVYRVKKFOFTSKHMEDESDLKEGGKRF 480
QY 481 GHICSHSPSCCTVNSNSWNCDEVLHSPAIEVRVHCOLVRLFARGIEE 529
DB 481 GKICKDRPSCTSASRSSWNCDEVMHSPAIEVRVHCOLVRLFARGIEE 529

RESULT 11

ADSI11054
ID ADSI11054 standard; protein; 536 AA.

AC ADSI11054;

DT 16-DEC-2004 (first entry)

DE Human therapeutic protein - SEQ ID 1291.

KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
KW aplastic anaemia; cancer; wound healing; gene therapy.

OS Homo sapiens.

FN WO2004080148-A2.

PD 23-SEP-2004.

PF 30-SEP-2003; 2003WO-US030720.

PR 02-OCT-2002; 2002US-0416186P.

PA (NUVE-) NUVELO INC.

PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;

XX WPI; 2004-668857/65.

DR N-PSDB; ADS10370.

XX New polynucleotide, useful in preparing a composition for diagnosing or
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT aplastic anemia or cancer for promoting wound healing.

PS Claim 20; SEQ ID NO 1291; 718pp; English.

XX The invention relates to a novel isolated polynucleotide and the encoded
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianaemic, cytostatic and vulnery activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during gene therapy
CC procedures. The current sequence is that of a human therapeutic protein
CC of the invention. The current sequence is not shown explicitly within the
CC specification but can be accessed from the WIPO web-site.

XX Sequence 536 AA;

Query Match 81.4%; Score 2350; DB 8; Length 536;

Best Local Similarity 78.1%; Pred. No. 1.6e-221;

Matches 452; Conservative 2; Mismatches 9; Indels 116; Gaps 4;

QY 32 LRWWR-----SPGPGAGAPGADACSPVSEIIAEEVDVHGKHQSGKWKMEKPYAFT 85

DB 1 MEWGRVFCVEMPAEGRGV-FTDACSPVSEIIAEEVDVHGKHQSGKWKMEKPYAFT 59

QY 86 VHCVRARRHRKWAQVFWCPPEQLCHLWLTQLREMLEKLTSPKHLVFINPFGKGQ 145

DB 60 VHCVRARRHRKWAQVFWCPPEQLCHLWLTQLREMLEKLT----- 100
QY 146 GKRIYERKVAFLFTLASITTDIIITEHANQAKETLYEINIDKYDGVCGDGMFSEVLH 205
DB 101 -----ITEHANQAKETLYEINIDKYDGVCGDGMFSEVLH 137
QY 206 GLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA----- 238
DB 138 GLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGFANDLAGDEVSVLSPVSGPEGGMV 197
QY 239 -----GSTDCVCYSTVGTSDAETSA 258
DB 198 HTFTLVLTALGCEHRSHPHFMDERTGTGHEVGFGENAGGLDGTDCVCYSTVGTSDAETSA 257
QY 259 LHIWVGDSLAMDVSVVHNSTLLRYSVSLGCGFYGDIIKDSEKKRWGLARYDFSGLKT 318
DB 258 LHIWVGDSLAMDVSVVHNSTLLRYSVSLGCGFYGDIIKDSEKKRWGLARYDFSGLKT 317
QY 319 FLSHHCEYGVTFPLPAQHTVGSPRDRKPCFVCRQSKQLEEEQKALYGLEAADV 378
DB 318 FLSHHCEYGVTFPLPAQHTVGSPRDRKPCFVCRQSKQLEEEQKALYGLEAADV 377
QY 379 BEWQVCGKFLAINATMNSCACRSPRGLSPAHLGDGSSDLILIRKCSRNFRLIRH 438
DB 378 BEWQVCGKFLAINATMNSCACRSPRGLSPAHLGDGSSDLILIRKCSRNFRLIRH 437
QY 439 TNQDQDFTFVEVYRVKKFOFTSKHMEDESDLKEGGKRFHICSSHPSCCTVNSNS 498
DB 438 TNQDQDFTFVEVYRVKKFOFTSKHMEDESDLKEGGKRFHICSSHPSCCTVNSNS 497
QY 499 WNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
DB 498 WNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 536

RESULT 12

ABR56301

ID ABR56301 standard; protein; 416 AA.

AC ABR56301;

DT 20-NOV-2003 (first entry)

DE Human Sphingosine kinase 4.

KW Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;
KW platelet transfusion; platelet stabiliser.

OS Homo sapiens.

PN WO2003031627-A1.

PD 17-APR-2003.

PF 28-SEP-2001; 2001WO-JP008537.

PR 28-SEP-2001; 2001WO-JP008537.

PA (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.

PI (CHBI-) CHEM BIOLOGY INST.

PI Igarashi Y, Kihara A;

XX WPI; 2003-354917/33.

DR N-PSDB; ACC70838.

XX Platelet derived polypeptides with sphingosine kinase activity for
PT treatment of sphingosine related disorders.

XX Claim 2; Page 29-30; 39pp; Japanese.

XX The present sequence is the protein sequence for human sphingosine kinase

CC 4 (SPHK4). The kinase can be used for the diagnosis and treatment of
CC sphingosine related disorders. The kinase can also be potentially used
CC for controlling toxicity of platelet transfusion and as a platelet
CC stabiliser

XX SQ Sequence 416 AA;

Query Match 76.5%; Score 2210; DB 6; Length 416;
Best Local Similarity 100.0%; Pred. No. 6.4e-208;
Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 122 MLEKLSRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLY 181
Db 1 MLEKLSRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLY 60
Qy 182 EINIDKVDGIVCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGST 241
Db 61 EINIDKVDGIVCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGST 120
Qy 242 DCVCYSTVGTSDAETLSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDSE 301
Db 121 DCVCYSTVGTSDAETLSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDSE 180
Qy 302 KRWGLGLARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRSQKQL 361
Db 181 KRWGLGLARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRSQKQL 240
Qy 362 EEOKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGSDLI 421
Db 241 EEOKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGSDLI 300
Qy 422 LIRKCSRFNRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKKEGKKRFG 481
Db 301 LIRKCSRFNRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKKEGKKRFG 360
Qy 482 HICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIBENPKPDSHS 537
Db 361 HICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIBENPKPDSHS 416

RESULT 13

ABG13541
ID ABG13541 standard; protein; 746 AA.

AC ABG13541;

DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #13532.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS77728.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 43900; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC patent did not appear in the printed specification. Note: The sequence data for this
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 746 AA;

Query Match 76.5%; Score 2208.5; DB 4; Length 746;
Best Local Similarity 75.8%; Pred. No. 2.1e-207;
Matches 430; Conservative 6; Mismatches 16; Indels 115; Gaps 6;

Qy 32 LRWR-----SQPGAGAGDADACVPVSEIIIAVETDVHGKHQSGKQKKEKPAFT 85
Db 200 MEWRRRVFCVEMFAEGRV-FTDACSVPVSEIIIAVETDVHGKHQSGKQKKEKPAFT 259
Qy 86 VHCVKARRHRHWKWAQVTFWCPEEQCLHMLQTLREMLEKLTSPKHLVFINPFGKQ 145
Db 259 VHCVKARRHRHWKWAQVTFWCPEEQCLHMLQTLREMLEKLTSPKHLVFINPFGKQ 299
Qy 146 GKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKYDGIIVCVGGDMFSEVLH 205
Db 300 -----ITEHANOAKETLYEINIDKYDGIIVCVGGDMFSEVLH 336
Qy 206 GLIGRTORSAGVDQNHPRVLPSSLRIGIIPA-----GSTDCVCYSTVGT 252
Db 337 GLIGRTORSAGVDQNHPRVLPSSLRIGIIPAHHVGPGENAGGLDSTDCVCYSTVGT 396
Qy 253 DAETSALHIV-----G 264
Db 397 DAETSALHIVVCCPEARKPPASRHTACGSGHGLCLGCLWNACLEASRLQSRMQSPG 456
Qy 265 DSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDEKKRWGLGLARYDFSLGKTLFLSHHC 324
Db 457 DSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDEKKRWGLGLARYDFSLGKTLFLSHHC 516
Qy 325 YEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRSQKQLEEBOKKALYGLEAAEDVEEQVV 384
Db 517 YEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRSQKQLEEBOKKALYGLEAAEDVEEQVV 576
Qy 385 CGKFLAINATNMSCACRRSPRGLSPAHLGDSGSDLIIRKCSRFNRLIRHTNQDQ 444
Db 577 CGKFLAINATNMSCACRRSPRGLSPAHLGDSGSDLIIRKCSRFNRLIRHTNQDQ 636
Qy 445 FDFTFVEVYRVKKQFTSKHMEDESDLKKEGKKRFGHICSSHPSCCTVSNSSWNCDE 504
Db 637 FDFTFVEVYRVKKQFTSKHMEDESDLKKEGKKRFGHICSSHPSCCTVSNSSWNCDE 696
Qy 505 VLHSPAIEVRVHCOLVRLFARGIBENP 531
Db 697 VLHSPAIEVR-----AVVMARGCKVP 718

RESULT 14	
ABG13543	
ID	ABG13543 standard; protein; 727 AA.
XX	
AC	ABG13543;
DT	18-FEB-2002 (first entry)
DE	Novel human diagnostic protein #13534.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
XX	
XX	(HYSE-) HYSEQ INC.
PA	
XX	
PI	Drmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	N-PSDB; AAS77730.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PS	Claim 20; SEQ ID NO 43902; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG030377 represent novel human diagnostic
CC	amino acid sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 727 AA;
Query Match	59.4%; Score 1714.5; DB 4; Length 727;
Best Local Similarity	63.7%; Pred No. 7, 4e-159;
Matches 345; Conservative	0; Mismatches 0; Indels 197; Gaps 3;
QY	170 TEHANOAKETLYEINIDKYDG-----190
Db	1 TEHANOAKETLYEINIDKYDGYTAHSRMPCCPRNHPAGKAAPVQGVGVGGQRAFPWEP 60
QY	191 -----IVCVGGDGMFSEVLHGL 207

Db	61	GPKQMPAKGVCRRLPLLLKLEASGLLRSEGRCTCRSPDRLCSSCSIVCVGGDGMFSEVLHGL	120
QY	208	IGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA-----	238
Db	121	IGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGPANDLAGDEVSVLSGVPVSGEGGMVHT	180
QY	239	-----GSTDVCVCYSTVGTSDAETSALH	260
Db	181	FTLVTLGCEHRSHPHFMDERRTGCEHVPGENAGGLDSTDCVCYSTVGTSDAETSALH	240
QY	261	IVV-----GDSLAMDVVS	272
Db	241	IVGCCPEARKPPASRHTACSGHGQLCLGQLNACLCEASRLQSRMQSPGDSLAMDVVS	300
QY	273	SVHENSTLLRYSVSLGFGFYGDIIKDSEKKRWLGRLARYDFSGLKTFLSHHCYEGTVSFL	332
Db	301	SVHENSTLLRYSVSLGFGFYGDIIKDSEKKRWLGRLARYDFSGLKTFLSHHCYEGTVSFL	360
QY	333	PAQHTVGSPPDRKPCRAGCFVCRQSKQOLEBEQKALYGLEAAEDVEWQVVCVKFLAIN	392
Db	361	PAQHTVGSPPDRKPCRAGCFVCRQSKQOLEBEQKALYGLEAAEDVEWQVVCVKFLAIN	420
QY	393	ATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNLFRLIRHTNQDQDFTFVEV	452
Db	421	ATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNLFRLIRHTNQDQDFTFVEV	480
QY	453	YRVKKFQFTSKHMEDESDLEKGGKRFHGHICSSHPSCCCTVSNSSMNCDEVLHSPAIE	512
Db	481	YRVKKFQFTSKHMEDESDLEKGGKRFHGHICSSHPSCCCTVSNSSMNCDEVLHSPAIE	540
QY	513	VR 514	
Db	541	VR 542	
	RESULT 15		
	ABB07854		
ID	ABB07854	standard; protein; 326 AA.	
XX			
AC	ABB07854;		
XX			
DT	03-JUL-2002	(first entry)	
XX			
DE		Human sphingosine kinase-like protein.	
XX			
KW		Human sphingosine kinase-like protein; intracellular signalling;	
KW		cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;	
KW		autoimmune disease; rheumatoid arthritis; Parkinson's disease.	
XX			
OS		Homo sapiens.	
XX			
PN		WO200228906-A2.	
XX			
PD		11-APR-2002.	
XX			
PF		05-OCT-2001; 2001WO-EP011516.	
XX			
PR		06-OCT-2000; 2000US-0238005P.	
XX			
PR		23-AUG-2001; 2001US-0314113P.	
XX			
PA		(FARB) BAYER AG.	
XX			
PI		Kossida S, Encinas J;	
XX			
DR		WPI; 2002-340094/37.	
DR		N-PSDB; ABL40822.	
XX			
PT		New reagent for modulating the activity of sphingosine kinase-like	
PT		protein polypeptide or polynucleotide and treating cancer, asthma,	
PT		allergy, an autoimmune disease, or a central or peripheral nervous system	
PT		disorder.	
XX			
PS		Claim 25; Fig 2; 120pp; English.	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 6, 2005, 02:35:14 ; Search time 4294.05 Seconds
(without alignments)
3678.672 Million cell updates/sec

Title: US-10-631-958-2

Perfect score: 1717

Sequence: 1 PKLLVFINPFGGKGQKRIKSRFNFLRLIRHTNQDQ 326

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10631958/runat.02092005.165815.4533/app.query.fasta_1.1941
-DB=GenEmbl -QFMT=fastcap -SUFFIX=rge -MINMATCH=0 -L-LOOPCL=0 -L-LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CGN 1.1.11756 @runat.02092005.165815.4533 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1717	100.0	979	6	AX456998 Sequence
2	1717	100.0	1840	6	AX224383 Sequence
3	1640.5	95.5	1654	9	CR456404 Homo sapi
4	1640.5	95.5	2042	9	HA457828 Homo sapi

5	1640.5	95.5	4171	6	BD183468	BD183468	Novel gen
6	1640.5	95.5	4171	9	AB051433	AB051433	Homo sapi
7	1640.5	95.5	4413	6	AX457006	AX457006	Sequence
8	1640.5	95.5	4445	9	AB079066	AB079066	Homo sapi
9	1633.5	95.1	4432	6	AR541900	AR541900	Sequence
10	1632.5	95.1	4463	6	BD102675	BD102675	Ceramide
11	1503.5	87.6	2830	10	AB079067	AB079067	Mus muscu
12	1503.5	87.6	3661	10	AK129416	AK129416	Mus muscu
13	1454	84.7	1459	6	CQ730476	CQ730476	Sequence
14	1452	84.6	1772	9	BC067255	BC067255	Homo sapi
15	1251.5	72.9	1450	5	CR386590	CR386590	Gallus ga
16	1140	66.4	1520	5	BC074110	BC074110	Xenopus l
17	1117.5	65.1	2494	5	BC074350	BC074350	Xenopus l
18	811	47.2	474	6	AX457001	AX457001	Sequence
19	650.5	37.9	2333	3	AK112750	AK112750	Ciona int
20	613	35.7	1810	3	AY061001	AY061001	Drosophil
21	613	35.7	2064	6	AR509346	AR509346	Sequence
22	613	35.7	2173	6	CQ608040	CQ608040	Sequence
23	613	35.7	3178	3	BT014925	BT014925	Drosophil
24	613	35.7	10337	6	CQ608039	CQ608039	Sequence
25	613	35.7	22353	2	AC017269	AC017269	Drosophil
26	613	35.7	181009	3	AC007575	AC007575	Drosophil
27	613	35.7	295289	3	AE003603	AE003603	Drosophil
28	566	33.0	1084	6	AR510195	AR510195	Sequence
29	508.5	29.6	329	6	AX457002	AX457002	Sequence
30	498.5	29.0	2257	8	AY362552	AY362552	Arabidops
31	473	27.5	2263	8	AY035126	AY035126	Arabidops
32	436	25.4	1599	9	AY357073	AY357073	Homo sapi
33	414.5	24.1	564	6	AX718996	AX718996	Sequence
34	389	22.7	1458	8	BT006621	BT006621	Arabidops
35	389	22.7	1757	8	AY128394	AY128394	Arabidops
36	389	22.7	70890	9	HSDA59H18	HSDA59H18	Human DNA
37	386	22.5	181772	2	AC007689	AC007689	Homo sapi
38	356.5	20.8	219077	2	AC116764	AC116764	Mus muscu
39	355	20.7	1799	8	AK072068	AK072068	Oryza sat
40	355	20.7	1817	8	AK107062	AK107062	Oryza sat
41	347	20.2	24083	2	AC095955	AC095955	Rattus no
42	347	20.2	260331	2	AK135400	AK135400	Rattus no
43	339.5	19.8	2451	9	AK129976	AK129976	Homo sapi
44	329	19.2	79048	2	AC008298	AC008298	Drosophil
45	319	18.6	32314	3	AF098993	AF098993	Caenorhab

ALIGNMENTS

RESULT 1	AX456998	Sequence 1 from Patent WO228906.	979 bp	DNA	linear	PAT 06-JUL-2002
LOCUS	AX456998	Sequence 1 from Patent WO228906.	979 bp	DNA	linear	PAT 06-JUL-2002
DEFINITION	AX456998	Sequence 1 from Patent WO228906.	979 bp	DNA	linear	PAT 06-JUL-2002
ACCESSION	AX456998	Sequence 1 from Patent WO228906.	979 bp	DNA	linear	PAT 06-JUL-2002
VERSION	AX456998.1	GI:21715789				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1					
AUTHORS	Kossida, S. and Encinas, J.					
TITLE	Regulation of human sphingosine kinase-like protein					
JOURNAL	Patent: WO 0228906-A 1 11-APR-2002;					
FEATURES	Bayer Aktiengesellschaft (DE)					
source	Location/Qualifiers					
	1..979					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					

ORIGIN

Alignment Scores:					
Pred. No.:	2.77e-153	Length:	979		
Score:	1717.00	Matches:	326		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		

```
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-631-958-2 (1-326) x AX456998 (1-979)

Qy 1 ProlysisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLeuArgIle 20
Db 2 CCAAAGCATTTACTCGTATTTATCAACCCGTTTGGAGGAAAAGGACAAGGCAAGCGGATA 61
Qy 21 TyrGluArgIleValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 62 TATGAAAGAAAAGTGGACCACTGTTACCTTAGCCTCCATCACCCTGACATCATCGGT 121
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 122 AACAAATCTATGTTAACTATGTAGAGTAACTTACTGAACATGCTAATCAGGCCCAAGGAG 181
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 182 ACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGT 241
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyValThrGlnArgSerAlaGlyValAsp 100
Db 242 ATGTTTCAGCAGGGTCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGTTCGAC 301
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 302 CAGAACCAACCCCGGGCTGTCTGGTCCCAAGTAGGCTCCGGATTTGGATCATTTCCGCGCA 361
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 362 GGGTCAACGGACTGCGTGTGTACTCCACCGTGGGACACGACGACGAGAACCTCGGCG 421
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 422 CTGCATATCTGTTGGGAGATCGCTGGCCATGATGTGTCCTCAGTCCACCAACAGCAGC 481
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 482 ACACCTCTTCCTACTCCGTCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 541
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 542 GACAGTGAAGAAAACGGTGGTGGGCTCTGCCAGATACGACTTTTCAGGTTTAAAGACC 601
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 602 TTCTCTCTCCACCACTGCTATGAAGGGACAGTGTCCTTCTCCCTGCAACAACACACCGTG 661
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 662 GGATCTCCAAGGATAGGAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCCAAGCAAG 721
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 722 CAGCAGCTGGAGGAGGAGCAGAGAACAACACTGTATGTTTGGAAAGCTGCGGAGGACGTG 781
Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 782 GAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAAATGTCTCTGT 841
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 842 GCTTGTCCCGGAGGCCCGAGGGCTCTCCCGGGCTGCGCACTTGGGAGACGGGTCTCT 901
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 902 GACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 961
Qy 321 ThrAsnGlnAspGln 326
Db 962 ACCAACCAACGACGGACCAAG 979
```

```
RESULT 2
AX224383 1840 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION Sequence 5 from Patent WO0160990.
ACCESSION AX224383
VERSION AX224383.1 GI:15554633
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Rastelli, L.
TITLE Novel sphingosine kinases
JOURNAL Patent: WO 0160990-A 5 23-AUG-2001;
Curagen Corporation (US); GENENTECH, INC. (US)
FEATURES
source 1. 1840
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 6.3e-153 Length: 1840
Pred. No.: 1717.00 Matches: 326
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-10-631-958-2 (1-326) x AX224383 (1-1840)
Qy 1 ProlysisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLeuArgIle 20
Db 265 CCAAAGCATTTACTCGTATTTATCAACCCGTTTGGAGGAAAAGGACAAGGCAAGCGGATA 324
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 325 TATGAAAGAAAAGTGGACCACTGTTCCACCTTAGCCTCCATCACCCTGACATCATCGGT 384
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 385 AACAAATCTATGTTAACTATGTAGAGTAACTTACTGAACATGCTAATCAGGCCCAAGGAG 444
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 445 ACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGT 504
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyValThrGlnArgSerAlaGlyValAsp 100
Db 505 ATGTTTCAGCAGGGTGTGTCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGTTCGAC 564
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 565 CAGAACCAACCCCGGGCTGTGCTGCCAGTAGCCTCCCGGATTTGGAAATCATTTCCGCGCA 624
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 625 GGGTCAACGGACTCGGTGTGTTACTCCACCGTGGGACCAACGACGACGAGAACCTCGGCG 684
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 685 CTGCATATCGTTGTTGGGACTCGCTGCCCATGGATGTGTCCTCAGTCCACCAACAGCAGC 744
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 745 ACACCTCTCGTACTCCGTGTCCTGTCGGCTACGGCTTCTACGGGGACATCATCAAG 804
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 805 GACAGTGAAGAAAACGGTGGTGGTCTTGCACAGATACGACTTTTCAGGTTTAAAGACC 864
```

QY 201 PheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
 Db 865 TTCTCTCCACACATGCTATGAAGGACAGTGTCTTCTCTCCCTGACAAACACACGGTG 924
 QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
 Db 925 GGATCTCCAAGGATAGGAAGCCCTGCGCGGAGGATGCTTTGTTGAGGCAAGCAAG 984
 QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
 Db 985 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGGACGTG 1044
 QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaAlaLeuAsnAlaThrAsnMetSerCys 280
 Db 1045 GAGGAGTGGCAAGTCTCTGTGGAGAGTTCTGGCCATCAATGCCACAAACATGTCCTGT 1104
 QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
 Db 1105 GCTTGTGCGCGAGGAGCCAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGCTTCT 1164
 QY 301 AspLeuLeuLeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuArgHis 320
 Db 1165 GACCTCATCTCATCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1224
 QY 321 ThrAsnGlnAspGln 326
 Db 1225 ACCAACCCAGCAGGACCAAG 1242

RESULT 3

CR456404 1654 bp mRNA linear PRI 25-MAY-2004
 LOCUS Homo sapiens BK29F11.1 full length open reading frame (ORF) cDNA
 DEFINITION clone (cDNA) C22ORF:pgem.bk29f11.1).
 ACCESSION CR456404
 VERSION CR456404.1 GI:47678338
 KEYWORDS CDNA; chromosome 22; ORF.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1654)
 AUTHORS Collins, J.E., Wright, C.L., Edwards, C.A., Davis, M.P., Grinham, J.A.,
 Cole, C.G., Goward, M.E., Aguado, B., Mallya, M., Mokrab, Y.,
 Huckle, E.J., Beare, D.M. and Dunham, I.
 TITLE Direct Submission
 JOURNAL Submitted (24-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: c22g@sanger.ac.uk Manuscript
 COMMENT Sanger Institute name : pgem.bk29f11.1
 Homo sapiens cDNA sequence. This sequence was generated as part of
 The Wellcome Trust Sanger Institute program to isolate cDNA clones
 representing the full length open reading frame of well annotated
 protein coding genes on human chromosome 22. For more information
 see http://www.sanger.ac.uk/HGP/Chr22/.

FEATURES

Location/Qualifiers
 1..1654
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="22"
 /clone="pgem.bk29f11.1"
 /lab_host="JM109"
 1..1654
 /gene="BK29F11.1"
 11..1624
 /gene="bk29f11.1"
 /codon_start=1
 /protein_id="CAG30290.1"
 /db_xref="GI:47678339"
 /translation="MGATGAAPLQSVLWVKQRCVSLSEPARALLRWSPGPGAGA
 PGADACSVPSVSIIVAEETDVHGKHQSGKWKMEKPAFTVHCVKRRARRRWKWAQV
 TFWCPPEQLCHLWLOTLEMLKLTSPKHLVFINPFGKGQGRIVYERKVALPFTL
 ASITTDIIUTEHANQAKETLYEINIDKVDGIVCVGGDMFSEVLHGLIGRTQRSAGVD
 QNHPRAVLVFSSLRIGIIPAGSTDCVSYSTVGTSDAETSAHLHIVVGDMLAMDVSSVHH

gene

CDS

ORIGIN

Alignment Scores:

Pred. No.: 9.96e-146 Length: 1654
 Score: 1640.50 Matches: 314
 Percent Similarity: 96.63% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 95.54% Indels: 11
 Db: 9 Gaps: 1

US-10-631-958-2 (1-326) x CR456404 (1-1654)

QY 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
 Db 398 CCAAGCAATTTACTGGTATTTATCAACCGTTTGGAGGAAAGCAAGGACAGCGGATA 457
 QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
 Db 458 TATGAAAGAAAGTGGCACCACCTGTTCACTTAGCTCCATCACCACCTGACATCATC--- 514
 QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
 Db 515 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 544
 QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
 Db 545 ACTCTGTATGAGATTAAACATAGACAAATACACGCGCATCGTCTGTGTGCGGGAGATGCT 604
 QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
 Db 605 ATGTTCCAGGAGTGTCTGCACGGTCTGATTGGAGGACGACAGAGGAGCGCGGGTGCAC 664
 QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
 Db 665 CAGAACCAACCCCGGGTGTGCTGTCGCCAGTAGTACCTCCGGATTGGGAATCATTCGCCCA 724
 QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
 Db 725 GGGTCAACGAGTCTGCTGTGTATTCTCCACCTGGGACGACGAGCGAGAACCTTCGGCG 784
 QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
 Db 785 CTGCATATCGTTGTTGGGACTCTCGTGGCCATGGATGTCTCTCAGTCCACCACCAACAGC 844
 QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
 Db 845 ACACCTCTTCGCTACTCCGTCCTGCTGCGGCTACGGCTTCTACGGGGACATCATCAAG 904
 QY 181 AspSerGluLysLysArgTyrLeuGluAlaArgTyrAspPheSerGlyLeuLysThr 200
 Db 905 GACATGTAGAAGAACCGTGTGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 964
 QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
 Db 965 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGACACACACACGCTG 1024
 QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
 Db 1025 GGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAGCAAG 1084
 QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
 Db 1085 CAGCAGCTGGAGGAGGAGCAGAGAAAGACACTGTATGTTTGGAGCTGCGGAGGACGTG 1144
 QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
 Db 1145 GAGGAGTGGCAAGTCTCTGTGGGAGTTTCTGGCCATCAATGTGCCACCAACATGTCTCTGT 1204

51
 43
 427
 1314
 320
 53


```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE       1 (bases 1 to 4171)
JOURNAL     Ohara,O., Nagase,T. and Nakajima,D.
            Novel genes and proteins encoded by the genes
            Patent: JP 2002345492-A 181 03-DEC-2002;
            KAZUSA DNA RESEARCH INSTITUTE
COMMENT     OS Homo sapiens (human)
            PN JP 2002345492-A/181
            PD 03-DEC-2002
            PF 26-FEB-2002 JP 2002049009
            PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
            PC C12N15/09, C07K14/47//A61K31/711.A61K38/00.A61K48/00.A61P25/00,
            A61P25/14,
            PC A61P25/18,A61P35/00,C12N15/00,A61K37/02
            CC Novel genes and proteins encoded by the genes FH Key
            Location/Qualifiers
            FT CDS
            Location/Qualifiers
            source
            1. .4171
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      3,32e-145      Length:      4171
Score:          1640.50      Matches:     314
Percent Similarity: 96.63%      Conservative: 1
Best Local Similarity: 96.32%      Mismatches: 0
Query Match:      95.54%      Indels:      11
DB:              6      Gaps:         1

US-10-631-958-2 (1-326) x BD183468 (1-4171)

QY 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 221 CCAAGCATTTACTGGTATTTATCAACCGTTTGGAGGAAAAGGCAAGGCAAGCGGATA 280

QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 281 TATGAAGAAAAGTGGCCACTGTGTACCTTAGCTTCATCCACCTGACATCATC--- 337

QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 338 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 367

QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysAspGly 80
Db 368 ACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTGTGTGGGGAGATGGT 427

QY 81 MetPheSerGluValIleuHisGlyLeuIleGlyArgThrGlnAtcSerAlaGlyValAsp 100
Db 428 ATGTTTCAGCGAGTGTGCACGGTCTGATTTGGAGAGCGACAGAGAGCGCGGGTCCAC 487

QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 488 CAGAACCCACCCCGGGCTGTGTGTCCTCCAGTAGCCTCCGGATTGGATCATTCCTCGCA 547

QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 548 GGGTCAACCGACTGGCGTGTGTACTCCACCGTGGGCACGACGACGACGAGAACTTCGGCG 607

QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisSer 160
Db 608 CTGCATATCGTTGTTGGGACATCGCTGGCCATGGATGTCTCTAGTCCACACACACAGC 667

QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 668 ACATCTCCTTCGTACTCTCCGTGTCCCTGTGGGCTAGGGCTTCTACGGGACATCATCAG 727

```

```

QY 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 728 GACAGTGAAGAAGAACCGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 787

QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 788 TTCTCTCTCCACCACTGCTATGAGGACAGTGTCTTCTCTCCCTGCACACACACACGGTG 847

QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 848 GGATCTCCAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTGCAGGCAAGCAAG 907

QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 908 CAGCAGCTGGAGGAGGAGCAGAGAAGACACTGTATGGTTTGGAGCTGCCGAGACCGTG 967

QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 968 GAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGCCCATCAATGCCACAAACATGTCCTGT 1027

QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1028 GCTTGTGCCGAGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTCT 1087

QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1088 GACCTCATCTCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCAC 1147

QY 321 ThrAsnGlnGlnAspGln 326
Db 1148 ACCAACCCAGCAGGACCAG 1165

RESULT 6
AB051433
LOCUS      AB051433
DEFINITION Homo sapiens mRNA for KIAA1646 protein, partial cds.
ACCESSION AB051433
VERSION    AB051433.1 GI:13359166
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS     Hirose,M., Nagase,T., Murahashi,Y., Kikuno,R. and Ohara,O.
TITLE       Identification of novel transcribed sequences on human chromosome
            22 by expressed sequence tag mapping
JOURNAL     DNA Res. 8 (1), 1-9 (2001)
MEDLINE     21156230
PUBMED      11258795
REFERENCE   2 (bases 1 to 4171)
AUTHORS     Ohara,O., Nagase,T. and Kikuno,R.
TITLE       Direct Submission
JOURNAL     Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,
            Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnaifookazusa.or.jp,
            URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
            Fax:81-438-52-3914)
FEATURES
            Location/Qualifiers
            source
            1. 4171
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="hk01650"
            /note="vector:pBluescriptII SK plus"
            1. 4171
            /gene="KIAA1646"
            <1. .1447
            /gene="KIAA1646"
            /note="Start codon is not identified."
            /codon_start=2
            /product="KIAA1646 protein"
            /protein_id="BAB33316.1"

```

/db_xref="GI:13359167"
 /translation="IIAVEETDVRHGKQSGKQWMEKPYAFTVHCVRARRHRWKA
 QVTWCPEQLCHLWLOTREMLKLTSRPKHLVFINPFGKGQKRIYVERKVAPLF
 TLASITTDLIIVTEHANAKETLYEINIDKYDGIVCVGDDGMFSEVLHGLIGRTORSAG
 VDNHPRALVVPSSIRGIIPAGSTDVCVSTVGTSDAETSALHIVGDSLAMDVSF
 HNSTLLRYSVSLIGYFGYDIIKDSSEKKRWLGARYDFSLKTLFLSHHCYEGTVSFL
 PAQHTVSGPRDKPCRCAGFCVCRQSKQLEEKALYGLEAAEDVEEWQVCGKFLA
 INATNMSCACRRSPRGLSPAHLGDSGLILIRKCSRFPNRLFLRIHTNQODQDFT
 FVEYRVKFOFTSKHMEDESDLKEGKKRFGHICSSHPSCCCTVSNSSWNCDEVL
 HSPAIEVRVHCQLVRLFARGIEENPKPDSHS"

ORIGIN

Alignment Scores:
 Pred. No.: 3.32e-145 Length: 4171
 Score: 1640.50 Matches: 314
 Percent Similarity: 96.63% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 95.54% Indels: 11
 DB: 9 Gaps: 1

US-10-631-958-2 (1-326) x AB051433 (1-4171)

Qy 1 ProlyshHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
 Db 221 CCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGCAAGCAAGCGGATA 280

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
 Db 281 TATGAAAGAAAGTGGACCACTGTTTACCTTAGCTCCATCACCATCATCATC--- 337

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
 Db 338 -----GTTACTGAACATGCTAATCAGGCCCAAGGAG 367

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyAspGly 80
 Db 368 ACTCTGTATGAGATTAAATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGT 427

Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
 Db 428 ATGTTTACGAGAGTGTCTACGGTCTGATTTGGAGGACGACAGAGCGCGGGGTGCAC 487

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
 Db 488 CAGAACCAACCCCGCGCTGTCTGTCGCCAGTAGCTCCGATTTGGAATCATTTCCCGCA 547

Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
 Db 548 GGGTCAACGGACTGCGTGTGTACTTCCACCGTGGCGCACACGACGACGAGAACCTCGCG 607

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
 Db 608 CTGCATATCGTTTGGGAGCTCGCTGGCCATGGATGTCTCTAGTCCACCAACCAACAGC 667

Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
 Db 668 ACACCTCTTCGCTACTCGGTGTCTGTCGGCTAGCGGCTTCTACGGCGGACATCATCAAG 727

Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
 Db 728 GACAGTGAGAGAAACCGTGGTGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 787

Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
 Db 788 TTCCTCTCCCACTGCTATGAAAGGACAGTGTCCTTCTCCCTGTCACACACACACGGTG 847

Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
 Db 848 GGATCTCCAAGGATAGGAAGCCCTGCGCGGCGAGGATGCTTTGTTTTCAGGCAAGCAAG 907

Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
 Db 908 CAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTGCGGAGGACGTG 967

Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
 Db 968 GAGGAGTGGCAAGTCGTCTGTGGAGAGTTTCTGCCCATCAATGCCACAAACATGCTCTGT 1027

Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
 Db 1028 GCTTGTCCGGAGCCGCCAGGGGCTCTCCCGGCTGCCACATTGGGAGACGGGTCTTCT 1087

Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
 Db 1088 GACCTCATCTCTATCCGGAATGCTCCAGGTTCATATTTCTGAGATTCTCATCAGGCAC 1147

Qy 321 ThrAsnGlnGlnAspGln 326
 Db 1148 ACCAACCCAGCAGGACCAG 1165

RESULT 7
 AX457006 4413 bp DNA linear PAT 06-JUL-2002
 LOCUS Sequence 9 from Patent WO0228906.
 DEFINITION AX457006
 ACCESSION AX457006
 VERSION AX457006.1 GI:21715795
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Kossida, S. and Encinas, J.
 TITLE Regulation of human sphingosine kinase-like protein
 JOURNAL Patent: WO 0228906-A 9 11-APR-2002;
 Bayer Aktiengesellschaft (DE)
 FEATURES
 Location/Qualifiers
 source 1..4413
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 76..1689
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD38256.1"
 /db_xref="GI:21715796"
 /translation="MGATGAAPLQSVLWVKQRCVSLPEFAPALLRWRSPGFGAGA
 PGADACSPVSEIIIAVEETDVRHGKQSGKQWMEKPYAFTVHCVRARRHRWKAQV
 TFWCPREQLCHLWLOTREMLKLTSRPKHLVFINPFGKGQKRIYVERKVAPLFTL
 ASITDIIIVTEHANAKETLYEINIDKYDGIVCVGDDGMFSEVLHGLIGRTORSAGVD
 NHPRALVVPSSIRGIIPAGSTDVCVSTVGTSDAETSALHIVGDSLAMDVSFVHH
 NSTLLRYSVSLIGYFGYDIIKDSSEKKRWLGARYDFSLKTLFLSHHCYEGTVSFLPA
 QHTVSGPRDKPCRCAGFCVCRQSKQLEEKALYGLEAAEDVEEWQVCGKFLA
 ATNMSCACRRSPRGLSPAHLGDSGLILIRKCSRFPNRLFLRIHTNQODQDFTFV
 EYRVKFOFTSKHMEDESDLKEGKKRFGHICSSHPSCCCTVSNSSWNCDEVLHS
 PAIEVRVHCQLVRLFARGIEENPKPDSHS"

ORIGIN

Alignment Scores:
 Pred. No.: 3.58e-145 Length: 4413
 Score: 1640.50 Matches: 314
 Percent Similarity: 96.63% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 95.54% Indels: 11
 DB: 6 Gaps: 1

US-10-631-958-2 (1-326) x AX457006 (1-4413)

Qy 1 ProlyshHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
 Db 463 CCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGCAAGCAAGCGGATA 522

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
 Db 523 TATGAAAGAAAGTGGACCACTGTTTACCTTAGCTCCATCACCATCATCATC--- 579


```
QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 580 -----GTTACTGAACTGCTAACTCAGGCCAAGGAG 609
QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 610 ACTCTGTATGAGATTAAACATACAGAAATACAGCGCATCTGTGTCTGGCGGAGATGTT 669
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 670 ATGTTTTCAGGAGGTGTCGACGCTCTGATTGGGAGACGACAGAGAGCGCGCGGTTCGAC 729
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 730 CAGAACCCCGCGGCTGTGCTGGTCCCATGAGCTCCGATTTGGATCATCTCCCGCA 789
QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 790 GGGTCAACGAGCTGCGTGTGTTACTCCACCGTGGGCACGACGACGAACTTCGGCG 849
QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 850 CTGCATATCTGTTTGGGAGCTCGCTGGCCATGGATGTCTCTCAGTCCACCACACAGC 909
QY 161 ThrLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 910 ACATCTCTTCGCTACTCCGCTGCTGCTGGCTACGGCTTCTACGGGACATCATCAAG 969
QY 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 970 GACAGTGAAGAAGACGCTGTTGGTCTTGCAGATACGACTTTTCAGTTTAAAGACC 1029
QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 1030 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTCTCTCCCTGCACACACACGCTG 1089
QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 1090 GGATCTCCAAAGGATAGGAAGCCCTCGCGGCGAGATGCTTTGTTGAGGCAAGCAAG 1149
QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 1150 CAGCAGCTGAGGAGGAGGACAGAAAGACACTGTATGTTTGAAGCTGCGGAGGACGTG 1209
QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 1210 GAGGAGTGGCAAGTCTGTCTGGGAAGTTTCTGGCCATCAATGCCACCAACATGCTCTGT 1269
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1270 GCTTGTGCGCGGAGCCCGAGGCGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 1329
QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1330 GACCTCATCTCTATCGGAAATGCTCCAGTTCAATTTTCTGAGATTCTCATCAGGCAC 1389
QY 321 ThrAsnGlnGlnAspGln 326
Db 1390 ACCAACCCAGCAGGACCAG 1407

RESULT 8
AB079066 4445 bp mRNA linear PRI 27-JUN-2002
LOCUS Homo sapiens cerk mRNA for ceramide kinase, complete cds.
DEFINITION
ACCESSION AB079066
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1.
AUTHORS Sugiura, M., Kono, K., Liu, H., Shimizugawa, T., Minekura, H.,
```

```
Spiegel, S. and Kohama, T.
Ceramide kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND
FUNCTIONAL CHARACTERIZATION
J. Biol. Chem. 277 (26), 23294-23300 (2002)
11956206
2 (bases 1 to 4445)
Sugiura, M., Kono, K., Shimizugawa, T., Minekura, H., Spiegel, S. and
Kohama, T.
Direct Submission
Submitted (29-JAN-2002) Masako Sugiura, Sankyo Co., Ltd.,
Pharmacology and Molecular Biology Research Laboratories; 2-58
Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan
(E-mail:mauglura@shina.sankyo.co.jp, Tel:81-3-3492-3131,
Fax:81-3-5436-8565)
Location/Qualifiers
1. .4445
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .4445
/gene="cerk"
124. .1737
/gene="cerk"
/codon_start=1
/product="ceramide kinase"
/protein_id="BAC01154.1"
/db_xref="GI:21624340"
/translation="MGATGAAPLQSVLMVKQRCVSLERAPALLRWRSPGPGQAGA
PGADACSVPSVEIIAVSETDVHGQSGKQMKPEYFTVCHVKRARRRHWKWAQV
TFWPEBQLCHLWQTLREMLEKLTSPKHLVFINPFGKQGRKIVYERKVALPFTL
ASITDIIVTEHANAKETLYEINIDKYGIVCVGGDMFSEVLHGLIGRTQRSAGVD
QNIIPRAVLPSLRIGIIPAGSTDCVCTSTVGTSDAETSALHI VVGDSLAMDVSSVHH
NTLLRYSVLLGYGYFDGDIKQSEKKRWGLARYDFSLGKTYFHSHCYEGTSLFPA
QHTYSPDRKPCFAGCVCRQKLEEQKALYGLAAEDVEEHWVCGVKFLAIN
ATNMSCARSPRLGSLPAHLGDSGLILIRKCSRFNPLRLRHTNQOQPDFTFV
EVYKVKFQTSKHMEDESDLKESGKKRFGHICSSHPSCCCTVSNSSMNCDEVLHS
PAIEVRVHCQLVRLPARGIBENPKPDSHS"
ORIGIN
Alignment Scores:
Pred. No.: 3.61e-145 Length: 4445
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 95.54% Indels: 11
DB: 9 Gaps: 1
US-10-631-958-2 (1-326) x AB079066 (1-4445)
QY 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 511 CCAAAGCATTTACTGTTATTTATCAACCCGTTTGAGGAAAGGACAAAGGCAAGCGGATA 570
QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 571 TATGAAAGAAAGTGGCACCACTGTTCACCTTAGCTCCATCACCACTGATCATC--- 627
QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 628 -----GTTACTGAACTGCTAACTCAGGCCAAGGAG 657
QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 658 ACTCTGTATGAGATTAAACATACAGAAATACAGCGCATCTGTGTCTGGCGGAGATGTT 717
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 718 ATGTTTTCAGGAGGTGTCGACGCTGTGATTGGGAGGACGACAGAGCGCGCGGTTCGAC 777
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 778 CAGAACCCCGCGGCTGTGCTGGTCCCATGAGCTCCCGATTTGGATTTGATTTCCCGCA 837
```

```
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 838 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACACGAGCAGAGAACCTCGGCG 897
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 898 CTGCATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTCTCAGTCCACCAACAGC 957
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleIys 180
Db 958 ACATCTCTTCCTACTCCGTGTCCCTGTCTGGGCTACGGCTTCTACGGGACATCATCAAG 1017
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 1018 GACAGTGAGAGAAACGGTGGTGGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 1077
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 1078 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTTCTCCCTGCACAAACACACGCTG 1137
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIys 240
Db 1138 GGATCTCAAGGGATAGGAAGCCCTGCCGGGCGAGATGCTTTGTTGAGGCAAGCAAG 1197
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 1198 CAGCAGCTGGAGGAGGACAGAGAAACACTGTATGGTTTGGAAAGCTGCGGAGGACGTG 1257
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 1258 GAGGAGTGGCAAGTCTGTCTGGGAAGTTTCTGGCCATCAATGCGCACAAACATGCTCTGT 1317
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1318 GCTTGTCCCGGAGGCCCGAGGGGCTCTCCCGGCTGCCCATCTGGGAGCGGGTCTTCT 1377
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1378 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCATCAGGCAC 1437
Qy 321 ThrAsnGlnInAppGln 326
Db 1438 ACCAACCCAGCAGGACCAG 1455

RESULT 9
AR541900 AR541900 4432 bp DNA linear PAT 08-OCT-2004
LOCUS
DEFINITION Sequence 148 from patent US 6743619.
ACCESSION AR541900
VERSION AR541900.1 GI:53933980
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 4432)
AUTHORS Tang, Y. T., Zhou, P., Goodrich, R., Liu, C., Asundi, V., Ren, F.,
Zhang, J., Zhao, Q. A., Yang, Y., Xue, A. J., Wehrman, T., Wang, J. -R.,
Wang, D., and Drmanac, R. T.
Nucleic acids and polypeptides
JOURNAL Patent: US 6743619-A 148 01-JUN-2004;
FEATURES Location/Qualifiers
source 1. 4432
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,66e-144 Length: 4432
Score: 1633.50 Matches: 313
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.01% Mismatches: 1
Query Match: 95.14% Indels: 11
DB: 6 Gaps: 1
```

```
US-10-631-958-2 (1-326) x AR541900 (1-4432)
Qy 1 ProlYshisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 1591 CCAAAGCATTTTACTGGTATTTTATCAACCCGTTTGGAGAAAGCAAGCAAGCGGATA 1650
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 1651 TATGAAAGAAAGTGGCACCACTGTTTCACTTACCTTACCTTCCATCACCCTGACATCATC- 1707
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 1708 -----GTTACTGAACATGCTAATCAGGCCAAGGAG 1737
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 1738 ACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTGTGTCTGGCGGAGATGGT 1797
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 1798 ATGTTCAAGCGAGGTGCTGCAAGCTCTGATTGGGAGGACGACAGAGGAGCGCGGGTCCAC 1857
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 1858 CAGAAACACCCCGGGCTGTGCTGCCAGTAGCTCCCGATTGGAAATCATTTCCCGCA 1917
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 1918 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACCAAGCAGCAGCAAAACCTCGGCG 1977
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 1978 CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACAGC 2037
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 2038 ACATCTCTCGTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAG 2097
Qy 181 AspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 2098 GACAGTGAGAAGAAACCGTGGTGGGCTTTGCCAGATACGACTTTTCAGGTTTAAAGACC 2157
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 2158 TTCTCTCTCCACCACTGCTATGGGAGGACAGTGTCTTCTTCTCTGACACACACCGGTG 2217
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 2218 GGATCTCCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAGCAAG 2277
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 2278 CAGCAGCTGGAGGAGGAGCAGAGAAACACTGTATGGTTTGGAAAGCTGGCGGAGGACGTG 2337
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 2338 GAGGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCGCACCAACATGCTCTGT 2397
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 2398 GCTTGTGCGCGGAGGCCCGCAGGGGCTCTCCCGGCTGCCCATCTGGGAGACGGGTCTTCT 2457
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 2458 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCATCAGGCAC 2517
Qy 321 ThrAsnGlnInAppGln 326
Db 2518 ACCAACCCAGCAGGACCAG 2535

RESULT 10
BD102675
```

LOCUS BD102675 4463 bp DNA linear PAT 27-AUG-2002
 DEFINITION Ceramide kinase and DNA thereof.
 ACCESSION BD102675
 VERSION BD102675.1 GI:22648249
 KEYWORDS WO 0196575-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 SUGIURA, M., KONO, K. and KOHAMA, T.
 Ceramide kinase and DNA thereof
 Patent: WO 0196575-A 1 20-DEC-2001;
 SANKYO CO LTD, MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA
 COMMENT OS Homo sapiens (human)
 PN WO 0196575-A/1
 PD 20-DEC-2001
 PF 11-JUN-2001 WO 2001JP004889
 PR 14-JUN-2000 JP OOP 178039
 PI MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA
 PC C12N15/54, C12N9/12, C12N1/21, C12Q1/48, C07K16/40, A61K31/7125, PC
 A61K48/00,
 PC A61P3/04, A61P3/10, A61P7/00, A61P25/00, A61P29/00, A61P31/18, PC
 A61P35/00
 CC Ceramide kinase and DNA thereof
 FH Key Location/Qualifiers
 FT CDS
 FEATURES
 source
 1. .4463
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,08e-144 Length: 4463
 Score: 1632.50 Matches: 312
 Percent Similarity: 96.32% Conservative: 2
 Best Local Similarity: 95.71% Mismatches: 1
 Query Match: 95.08% Indels: 11
 DB: Gaps: 1
 US-10-631-958-2 (1-326) x BD102675 (1-4463)
 QY 1 ProLysHisLeuLeuValPheIleLeuProPheGlyGlyGlyGlnGlyLysArgTle 20
 DB 511 CCAAAGCATTACTGCTATTTATCAACCGTTTGGAGGAAAGGACAAAGCGGATA 570
 QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
 DB 571 TATGAAGAAGAAAGTGGCACCCTGTCACCTTAGCCTCCATCACCACATCATC--- 627
 QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
 DB 628 -----GTTACTGACATGCTTAATCAGGCCAAGGAG 657
 QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
 DB 658 ACTCTGTATGAGATTAAACATACACAAATACGACGCATCGTCTGTGTGGCGGAGATG 717
 QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyIleValArgThrGlnArgSerAlaGlyValAsp 100
 DB 718 ATGTTTCAGCGAGTGTCTGCACGGTCTGATTGGGAGACGACAGAGAGCGCGGGTGCAC 777
 QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
 DB 778 CAGAACCAACCCCGGCTGTGTGTCCTCCAGTAGCCTCCGATTGGAATCATCCCGCA 837
 QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
 DB 838 GGTCAACAGCATGCTGTGTGTACTCCACCGTGGGCACGACGACGAGAACTCGGG 897
 QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisAsnSer 160

DB 898 CTGCATATCGTTGTTGGGACTCGCTGCCATGATGTGCTCTAGTCCACACAAACAGC 957
 QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
 DB 958 ACACCTCTTCGCTACTCGCTGCTCCCTGCTGGCTAGCGCTTCTACGGGACATCATCAAG 1017
 QY 181 AspSerGluLysLysArgTyrPheLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
 DB 1018 GACAGTGGAGAAGAACCGTGTGGTCTTCCACGATACGACTTTTTCAGGTTTAAAGACC 1077
 QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
 DB 1078 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTCCCTGCACACACACGGTG 1137
 QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
 DB 1138 GGATCTCAAGGATAGGAAGCCCTCCGCGGCGAGATGCTTTGTTGTCAGGCAAGCAAG 1197
 QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
 DB 1198 CAGCAGCTGGAGGAGGAGCAGAGAAGAACACTGTATGGTTTGGAGCTGCGGAGACGTG 1257
 QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
 DB 1258 GAGGAGTGGCAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTGT 1317
 QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
 DB 1318 GCTTGTGCGGAGGAGCCCGGCGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 1377
 QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
 DB 1378 GACCTCATCTCTCATCCGGAATGCTCCAAAGTTCATTTTCTGAGATTTCTCATCAGGCAC 1437
 QY 321 ThrAsnGlnGlnAspGln 326
 DB 1438 ACCAACACGACGAGGACCAG 1455
 RESULT 11
 AB079067
 LOCUS AB079067 2830 bp mRNA linear ROD 27-JUN-2002
 DEFINITION Mus musculus cerk mRNA for ceramide kinases, complete cds.
 ACCESSION AB079067
 VERSION AB079067.1 GI:21624341
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 SUGIURA, M., KONO, K., LIU, H., SHIMIZUGAWA, T., MINEKURA, H.,
 SPIEGEL, S. and KOHAMA, T.
 Ceramide Kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND
 FUNCTIONAL CHARACTERIZATION
 J. Biol. Chem. 277 (26), 23294-23300 (2002)
 REFERENCE 11956206
 2 (bases 1 to 2830)
 SUGIURA, M., KONO, K., SHIMIZUGAWA, T., MINEKURA, H., SPIEGEL, S. and
 KOHAMA, T.
 Direct Submission
 Submitted (29-JAN-2002) Masako Sugiura, Sankyo Co., Ltd.,
 Pharmacology and Molecular Biology Research Laboratories; 2-58
 Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan
 (E-mail: meugliura@shina.sankyo.co.jp, Tel: 81-3-3492-3131,
 Fax: 81-3-5436-8565)
 FEATURES
 source
 1. .2830
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1. .2830
 /gene="cerk"
 gene

/protein_id="BAC98226.1"
 /db_xref="GI:37360496"
 /translation="FCFAASRPKHLVLFNPPFGKQKRIYKTVAPLFTLASITTE
 IIVTHANQAKETLYEINTDYGIVCGDGMFSEVLHGVIQRTQOAGIDPNHPRA
 VLVSTLRIGIIPAGSTDCVYSTVGTNDATSAHLIIIGDSLADIVSVHVNLLR
 YVSLVGFYGDLLKDSKKRWMLGRVDFSLKATFLSHOYEGTFLFLPAQHTVGS
 PRDNKPCAGCFVCRSQKQLEEEKKALYGLNAAEEVEEMQVTCGKFLANATMNSC
 ACPRSPGGLSPAHILGDSDDLILIRKSRFNFLRLIRHTNQEQDFPTFVYVRVK
 KPHFTSKHVEDSDNDSKQEKQKFKICKDRPSCTCSASRSWNDCGVMHSPALIEV
 VHCQLVRLPARGIIEES"

ORIGIN

Alignment Scores:
 Pred. No.: 2,8e-132 Length: 3661
 Score: 1503.50 Matches: 283
 Percent Similarity: 92.64% Conservative: 19
 Best Local Similarity: 86.81% Mismatches: 13
 Query Match: 87.57% Indels: 11
 DB: 10 Gaps: 1

US-10-631-958-2 (1-326) x AK129416 (1-3661)

QY 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgile 20
 DB 23 CCGAAGCACTTCTGCTATTTCATCAACCTTTCGGAGGGAAGGTCAGGGCAAGCGCATC 82
 QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleleGly 40
 DB 83 TATGAAAAACAGTGGCGCTCTGTTTACCTTGGCTTCCATCACTACGAGATCATC--- 139
 QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
 DB 140 -----ATTACAGAGCATGCCAACCAAGCCAAAGGAG 169
 QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
 DB 170 ACTTATACAGATCAACACAGACAGCATATGATGTCATCTGTGTGTGGAGCGGC 229
 QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
 DB 230 ATGTTCCAGCAGGTCTGTCATGGGTGATTGGGAGGACGACAGAGCGCTGTTATCGAC 289
 QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
 DB 290 CCAATACACCCCGAGCGGTCTGTGTCAGTACCTCAGATCGGCATCATACCCGCA 349
 QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
 DB 350 GGGTCCACAGATTGTGTGTACTCAACAGTGGGCAACAGCAGCAGACATCGGCT 409
 QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSer 160
 DB 410 TTGCACATCATTTATGGGACTCGCTGGCAATAGACGTCTCTGTGCACTTACCATAAC 469
 QY 161 ThrLeuLeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
 DB 470 ACGTGTCTGCGTACTCGGTCTCTGTGTCGGCTACGGTTTCTACGGGACTTAATCAAG 529
 QY 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
 DB 530 GACAGGAAAGAACCGTGGATGGCCCTCGTCCGGTATGATTTCTCAGGTTGAAGACC 589
 QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
 DB 590 TTTCTCTCTCATCAGTACTATGAAGGACACTGTCTCTCTCCAGCAGCAGCAGCGGTG 649
 QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
 DB 650 GGATCTCCACGGGACAATAAACCTCGCGGCTGGGTGCTTCTGTGCGGAGCAGACGAG 709
 QY 241 GlnGlnLeuGluGluGlnLysLysValAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
 DB 710 CAACAGCTGGAAGAAGAGAGAAAGCCCTGTATGGCTCGAGAACCCCGGAGGAAGTG 769

QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
 DB 770 GAAGAGTGGCAAGTGACATGTGGGAAGTTCTCTGGCCATCAATGCCACCAACATGTCTGT 829
 QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
 DB 830 GCTTGTCTCGAGCCCTGGGGGCTGTCCCATTTGCCCATCTCGGAGATGGGTCTTCT 889
 QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
 DB 890 GACCTCATCTTATCCGGAAGTGTCTCCAGGTTCACTTCTCTGAGATTCTCATCCGGCAC 949
 QY 321 ThrAsnGlnGlnAspGln 326
 DB 950 ACGAACCAGGAGGACCAG 967
 RESULT 13
 LOCUS CQ730476 1459 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 16410 from Patent WO02068579.
 ACCESSION CQ730476
 VERSION CQ730476.1 GI:42304409
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 16410 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES
 source
 1..1459
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,21e-128 Length: 1459
 Score: 1454.00 Matches: 275
 Percent Similarity: 99.64% Conservative: 2
 Best Local Similarity: 98.92% Mismatches: 1
 Query Match: 84.68% Indels: 0
 DB: 6 Gaps: 0
 US-10-631-958-2 (1-326) x CQ730476 (1-1459)
 QY 49 GluValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAsp 68
 DB 344 AAGCTGATTACTGAACATGCTTAATCAGCCAGAGAGACTCTGTATGAGATTAACTATGAC 403
 QY 69 LysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGly 88
 DB 404 AAATACGACGCATCTGTCTGTCTGGCGAGATGTTATGTTTCAGCGAGTGTCTGACCGT 463
 QY 89 LeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeu 108
 DB 464 CTGATTGGAGAGCGCAGAGAGCGCGGGTTCGACCAAGAACCCACCCCGGGCTGTGCTG 523
 QY 109 ValProSerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyr 128
 DB 524 GTCCCCAGTACCTTCCGATTGGAATCATTTCCCGAGGGTCAACGAGCTGCTGTGTTTAC 583
 QY 129 SerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSer 148
 DB 584 TCCACCGTGGGACACGAGCAGCAACCTCGGGCTGCATATCGTTGTTGGGACTCG 643
 QY 149 LeuAlaMetAspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSer 168

Db	644	CTGGCCATGGATGTCTCAGTCACCAACACGACACACTCCTTCCTACTCCGTGTCC	703		
Qy	169	LeuLeuGlyTyrGlyPheTyrGlyAspIleIleIysAspSerGluIleYsLysArGTrpLeu	188		
Db	704	CTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGGTTT	763		
Qy	189	GlyLeuAlaArgTyrAspPheSerGlyLeuIysThrPheLeuSerHisHisCysTyrGlu	208		
Db	764	GGTCTTGGCAGATAGCACTTTTTCAGGTTTAAGACCTTCCTCTCCACCACTGCTATGAA	823		
Qy	209	GlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro	228		
Db	824	GGGACAGTGTCTCTCCTCCCTGCACACACACGCTGGGATCTCCAAGGATAGGAAGCCC	883		
Qy	229	CysArgAlaGlyCysPheValCysArgGlnSerIysGlnGlnLeuGluGluGlnLys	248		
Db	884	TGCCGGGAGGATGCTTTGTTTGCAGGCAACAGCAGCAGCTGGAGGAGGACAGAAG	943		
Qy	249	LysAlaLeuTyrGlyLeuGluAlaGluAspValGluGluTyrGlnValValCysGly	268		
Db	944	AAAGCACTGTATGTTTGGAGCTGCGGAGACGTGGAGAGTGGCAAGTCTGTGTGGG	1003		
Qy	269	LysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly	288		
Db	1004	AAGTTTCTGGCCATCAATGCCACAAACATGCTCTGTCTGTCTGCGCGAGCCCGAGGGC	1063		
Qy	289	LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgIysCys	308		
Db	1064	CTCTCCCGGCTGCCCATCTGGGAGACGGGTCTTCTGACCTCATCCTCCGGAATGC	1123		
Qy	309	SerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGln	326		
Db	1124	TCCAGTTTCATTTTCTGAGATTTCTCATCAGGCACACCAACACGAGGACCAG	1177		
RESULT 14	BC067255	1772 bp	mrna	linear	PRI 06-APR-2004
LOCUS	BC067255	1772 bp	mrna	linear	PRI 06-APR-2004
DEFINITION	Homo sapiens ceramide kinase, mRNA (cdna clone IMAGE:6185601), complete cds.				
ACCESSION	BC067255				
VERSION	BC067255.1 GI:45595582				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 1772) Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schreitz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.N., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
REFERENCE	2 (bases 1 to 1772)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-MAR-2004) National Institutes of Health, Mammalian				
Qy	1	ProlYsHisLeuLeuValPheIleAsnProPheGlyGlyLys-GlyGlnGlyLysArgIle	20		
Db	447	CCAAAGCATTTTACTGGTATTATCAAC-----AAAAGGACAAGGCGGAT	494		
REMARK COMMENT					
Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
NIH-MGC Project URL: http://mgc.nci.nih.gov					
Contact: MGC help desk					
Email: gcapsb-r@mail.nih.gov					
Tissue Procurement: Dr. James R. Lupski					
cDNA Library Preparation: Life Technologies, Inc.					
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada					
info@bcgsc.bc.ca					
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smalilus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.					
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov					
Series: IRAC Plate: 141 Row: d Column: 4					
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21703365					
This clone has the following problem: The cds is short compared to the longest cds in the locus.					
Location/Qualifiers					
1. 1772					
/organism="Homo sapiens"					
/mol_type="mRNA"					
/db_xref="taxon:9606"					
/clone="IMAGE:6185601"					
/tissue_type="Peripheral Nervous System, dorsal root ganglion"					
/clone_lib="Lupski_dorsal_root_ganglion"					
/lab_host="DH10B"					
/note="Vector: pCMV-SPORT6"					
1. 1772					
/gene="CERK"					
/note="synonyms: KIAA1646, hCERK, dA59H18.3, DKFP434E0211, FLJ21430, FLJ23239, dA59H18.2, LK4"					
/db_xref="LocusID:64781"					
579. 1598					
/gene="CERK"					
/codon_start=1					
/product="CERK protein"					
/protein_id="AAH67255.1"					
/db_xref="GI:45595583"					
/db_xref="LocusID:64781"					
/translation="MFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPAGSTD CVCYSTVGTSDAETSAIHIVGDSLAMDVSVDHNSITLLYSVSLGCGVGVGDIKDS EKKRUGLARYDFSGKLTFLFHCYEGTGVFLPQAHTVGSFPRKKPCRCAGCFVCRQSK QQLBEEQKALYGLEAAEDVEQVCGKFLAINATNMSCACRRSPRLGSPAHLGG SLDLILRKCSRFNRLFLIRTNQQDQDFTFVEYVRVKKFQFTSKHMEDESDLKE GKKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPALIEVRVHCOLVRLFARGIEENPK PDSHS"					
ORIGIN					
Alignment Scores:					
Pred. No.:					
Score:					
Percent Similarity:					
Best Local Similarity:					
Query Match:					
DB:					
US-10-631-958-2 (1-326) x BC067255 (1-1772)					

Db 570 CTTTCATATTATTATAGGTGACTGTGACGCTTTAGATGTTCTCATCTGTGTCATCAAAACAAC 629
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyVaspileleLys 180
Db 630 ACGTTTTGAAGTATGCTGTATCATTTGTGGCTATGTTTTTATGGAGATGTTTGAAA 689
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 690 GACAGTGAAGAAACGGTGGATGGTCCGATGAGATACGACTATTTCAGGCTTCAAGACT 749
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 750 TTTCTTTCTCATCATTTACTATCAAGGAACAATTTCTTTCAACCAGCAAAACACACACTG 809
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 810 GGATCTCCACGAGATAAGATAGCTGCAGACAGGATGTTATATTTGCAAGGAAGTGAG 869
Qy 241 GlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 870 CAACAACCTGGCAGAACACCGCAAGAAGTGT-----GGATTCAAAACATGAAGAAGATGAA 923
Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 924 GAAGAATCGAAGGTATTATAAGGGAAATTTCTAGCCATCAATGTAGTAAATATGTGCTGT 983
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyVaspileleSer 300
Db 984 GCCTGTCCACGAAGTCCAAAGGTCCTTCACGGCAGCTCATTTAGCAGATGGTTCAGCT 1043
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1044 GACCTGATTCAGTTTCGGAAATGCTCCAGATTTGATTTTCTGCGTTATCTTGTTCAGACAT 1103
Qy 321 ThrAsnGlnGlnAspGln 326
Db 1104 ACAACAAGAGATGACCAG 1121

Search completed: September 6, 2005, 16:37:43
Job time : 4353.05 secs

Run on: September 5, 2005, 15:12:23 ; Search time 4448.35 Seconds
(without alignments)
10664.096 Million cell updates/

Sequence: 1 accaaagcatttactggtat.....acaccaaccagcagqaccacq 979

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs. 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

[illegible]

Post-processing: Minimum Match 0%
Maximum Match 100%

Булганов С. С.

Database : GENEMOT; 1: qb ba:

```
2:  gb_inj:
3:  gb_in:*
```

```
3: gb_ov: *
5: gb_ov: *
```

7. 2. 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 8

ॐ नमो भगवते वासुदेवाय

11: 06 855: *

```
12:  qb_un: *
13:  qb_un: *
```

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1124: NO: IS THE NUMBER OF REPEATS
score greater than or equal to the

[illegible]

2
3
4
4
4
4
4
4
5
5
6

| Query | Score | Match | Length | DB | ID |
|-------|-------|-------|--------|----|----|
| No. | | | | | |

| | | | | | |
|---|-----|-------|-----|---|----------|
| 1 | 979 | 100.0 | 979 | 6 | AX456998 |
|---|-----|-------|-----|---|----------|

| | | | | | |
|---|-----|------|------|---|----------|
| 3 | 903 | 92.2 | 1654 | 9 | CR456404 |
|---|-----|------|------|---|----------|

| | | | | | |
|---|-----|------|------|---|----------|
| 5 | 903 | 92.2 | 4171 | 6 | BD183468 |
|---|-----|------|------|---|----------|

| | | | | | |
|---|-----|------|------|---|----------|
| 7 | 903 | 92.2 | 4413 | 6 | AX457006 |
|---|-----|------|------|---|----------|

| | | | | | |
|---|-------|------|------|---|----------|
| 9 | 901.4 | 92.1 | 4432 | 6 | AR541900 |
|---|-------|------|------|---|----------|

| | | | | | |
|----|-------|------|------|---|----------|
| 11 | 826.8 | 84.5 | 1459 | 6 | CQ730476 |
|----|-------|------|------|---|----------|

| | | | | | |
|----|-----|------|------|----|----------|
| 13 | 639 | 65.3 | 3661 | 10 | AK129416 |
|----|-----|------|------|----|----------|

| Run | Time | Temp | Pressure | Flow | Sample |
|-----|-------|------|----------|------|----------|
| 15 | 464.4 | 47.4 | 474 | 6 | AX457001 |
| 15 | 464.4 | 47.4 | 474 | 6 | AX457001 |

[illegible]

| 10 | 302.1 | 30.0 | 1320 | 3 | 20074110 | 2007411 |
|----|-------|------|------|---|----------|---------|
| 19 | 302.2 | 30.9 | 329 | 6 | AX457002 | AX45700 |

| | | | | | |
|------|-------|------|--------|---|----------|
| C 20 | 209.6 | 21.4 | 70850 | 3 | nsd43549 |
| C 21 | 209.8 | 21.4 | 70872 | 2 | AC007689 |
| C 22 | 200.6 | 20.5 | 564 | 6 | AX139956 |
| C 23 | 167.2 | 17.1 | 240483 | 3 | AC035955 |
| C 24 | 167.2 | 17.1 | 260331 | 2 | AC135400 |
| C 25 | 167 | 17.1 | 1637 | 6 | AX457003 |
| C 26 | 164 | 16.8 | 219077 | 2 | AC116764 |
| C 27 | 153 | 15.6 | 153 | 6 | AX457004 |
| C 28 | 152.4 | 15.6 | 181772 | 2 | AC007689 |
| C 29 | 136.8 | 14.0 | 532 | 6 | AX224385 |
| C 30 | 106.2 | 10.8 | 325530 | 3 | AC125699 |
| C 31 | 105.6 | 10.8 | 153756 | 5 | BX248240 |
| C 32 | 91.8 | 9.4 | 2333 | 3 | AK112703 |
| C 33 | 90.4 | 9.2 | 1323 | 5 | CR760136 |
| C 34 | 90.4 | 9.2 | 163647 | 2 | AC147916 |
| C 35 | 90.4 | 9.2 | 164496 | 2 | AC147881 |
| C 36 | 89 | 9.1 | 550 | 6 | AX457005 |
| C 37 | 83 | 8.5 | 267546 | 2 | AC125740 |
| C 38 | 78.6 | 8.0 | 1084 | 6 | AR510195 |
| C 39 | 78.6 | 8.0 | 3178 | 3 | BT014925 |
| C 40 | 77 | 7.9 | 1810 | 3 | AY061001 |
| C 41 | 77 | 7.9 | 2064 | 6 | AR509346 |
| C 42 | 77 | 7.9 | 2173 | 6 | CQ608040 |
| C 43 | 77 | 7.9 | 10337 | 6 | CQ608039 |
| C 44 | 77 | 7.9 | 22353 | 2 | AC017529 |
| C 45 | 77 | 7.9 | 181009 | 3 | AC007575 |

ALIGNMENTS

| | |
|------------|---------------------------------------------------------------------------------------------------|
| RESULT 1 | |
| LOCUS | AX456998 |
| DEFINITION | Sequence 1 from Patent WO0228906. |
| ACCESSION | AX456998 |
| VERSION | AX456998.1 GI:21715789 |
| KEYWORDS | 'Homo sapiens (human) |
| SOURCE | Homo sapiens |
| ORGANISM | Eukaryota; Metazoa; Chordata; Crata
Mammalia; Eutheria; Primates; Cat
1 |
| REFERENCE | Kossida,S. and Encinas,J.
Regulation of human sphingosine k
TITLE |
| JOURNAL | Patent: WO 0228906-A 11-APR-200
Bayer Aktiengesellschaft (DE)
Location/Qualifiers
1..979 |
| FEATURES | source |
| | /organism="Homo sapiens"
/mol_type="unassigned DN
/db_xref="taxon:9606" |

| ORIGIN | Query Match | 100.0%; Score 979; |
|--------|---------------------------|--------------------------------------|
| | Best Local Similarity | 100.0%; Pred. No. 3. |
| | Matches 979; Conservative | 0; Mismatchatches |
| QY | 1 | ACCAAGCATTTACTGGTATTTTATCAACCCGTTT |
| DB | 1 | ACCAAGCATTTACTGGTATTTTATCAACCCGTTT |
| QY | 61 | ATATGAAAGAAAGTGGCACCACCTGTTCAACCTTT |
| DB | 61 | ATATGAAAGAAAGTGGCACCACCTGTTCAACCTTT |
| QY | 121 | TACCAATTTCTATGTTAACTATGTAGAAGTAAT |
| DB | 121 | TACCAATTTCTATGTTAACTATGTAGAAGTAAT |
| QY | 181 | GACTCTGTATGAGATTAA CATAGACAAATATACGA |
| DB | 181 | GACTCTGTATGAGATTAA CATAGACAAATATACGA |

[illegible]

| | | | | |
|-------------------------------------------------------------------|------|------------------------------------------------------------|------|---------------------------------------------------------------|
| QY | 361 | AGGGTCAACGAGCTGCGTGTGTTACTCCACCGTGGGACACGACGCGAGAAACCTCGGC | 420 | PGADACSVPVSEIIIAVEETDVHGKHGSGKQWMEKPYAFTVHCVKRARRHRWKAQV |
| Db | 547 | AGGGTCAACGAGCTGCGTGTGTTACTCCACCGTGGGACACGACGCGAGAAACCTCGGC | 606 | TFWCPBOLCHLMLOTLREMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTL |
| QY | 421 | GCTGCATATCGTTGTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAG | 480 | ASITBDILVTEHANQAKELYEINIDKYDGIVCVGGDMFSEVLHGLIGLQTSQSGVD |
| Db | 607 | GCTGCATATCGTTGTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAG | 666 | QNTPRAVLVPESSURIGIIPAGSIDCVCTIVTSGDASTALHIVVGLDMLANDVSSVHH |
| QY | 481 | CACACTCTTCGCTACTCGTGTCCCTGCTGGGTACGGCTTCTACGGGACATCATCA | 540 | NSTLLRYSVLLGFGYGDIIKDSEKKRWGLARYDFSLKTLFUSHHCYBGTYSFLPA |
| Db | 667 | CACACTCTTCGCTACTCGTGTCCCTGCTGGGTACGGCTTCTACGGGACATCATCA | 726 | QHTVGSFPRDKPRACGFCVQRQQLLEEQKALYGLAAEDVEEWQVCGKFLAIN |
| QY | 541 | GGACAGTCAGAAAGAACCGTGTGGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGAC | 600 | ATNWSACACRRSPRGLSPAHLGDGSDLIIRKCSRNFRLRIHTNQDQDFTFV |
| Db | 727 | GGACAGTCAGAAAGAACCGTGTGGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGAC | 786 | EVTVKFKOFTSKHMEDESDLKEGGKKRGHICSSHPSCCTTVSNSSWNCNDCGVLHS |
| QY | 601 | CTTCCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCGCTGCACAAACACGGT | 660 | PAIEVRVHCOLVRLFARGIBENPKPDSH* |
| Db | 787 | CTTCCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCGCTGCACAAACACGGT | 846 | ORIGIN |
| QY | 661 | GGGATCTCCAGGATAGAGACCTCGCGGGCAGGATGCTTTGTTGCAGGCAAGCAA | 720 | Query Match 92.2%; Score 903; DB 6; Length 4413; |
| Db | 847 | GGGATCTCCAGGATAGAGACCTCGCGGGCAGGATGCTTTGTTGCAGGCAAGCAA | 906 | Best Local Similarity 96.6%; Pred. No. 6.9e-229; |
| QY | 721 | CGACGAGCTGGAGGAGCAGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGT | 780 | Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1; |
| Db | 907 | CGACGAGCTGGAGGAGCAGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGT | 966 | |
| QY | 781 | GGAGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCAAACATGTCTTG | 840 | |
| Db | 967 | GGAGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCAAACATGTCTTG | 1026 | |
| QY | 841 | TGCTTGTCCGCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGAGACGGGTCTTC | 900 | |
| Db | 1027 | TGCTTGTCCGCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGAGACGGGTCTTC | 1086 | |
| QY | 901 | TGACCTCATCTCATCCGGAATGCTCCAGGTTCAAATTTCTGAGATTTCTCATCAGGCA | 960 | |
| Db | 1087 | TGACCTCATCTCATCCGGAATGCTCCAGGTTCAAATTTCTGAGATTTCTCATCAGGCA | 1146 | |
| QY | 961 | CACCAACAGCAGGACACAG 979 | | |
| Db | 1147 | CACCAACAGCAGGACACAG 1165 | | |
| RESULT 7 | | | | |
| AX457006 | | | | |
| LOCUS | | | | |
| DEFINITION | | | | |
| ACCESSION | | | | |
| VERSION | | | | |
| KEYWORDS | | | | |
| SOURCE | | | | |
| ORGANISM | | | | |
| Homo sapiens (human) | | | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| 1 | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| Bayer Aktiengesellschaft (DE) | | | | |
| FEATURES | | | | |
| source | | | | |
| 1..4413 | | | | |
| /organism="Homo sapiens" | | | | |
| /mol_type="unassigned DNA" | | | | |
| /db_xref="taxon:9606" | | | | |
| 76..1689 | | | | |
| /note="unnamed protein product" | | | | |
| /codon_start=1 | | | | |
| /protein_id="CAD38256.1" | | | | |
| /db_xref="GI:21715796" | | | | |
| /translation="MGATGAEEPLQSVLWVKQRCVSLPAPALLRWRSPPGAGA | | | | |

```
Db 1269 TGCTTGTCCGCGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGAGGGGTCTTC 1328
Qy 901 TGACCTCATCTTCATCCGGAATGCTCCAGGTTCATTTCTGAGATTCTCATCAGGCA 960
Db 1329 TGACCTCATCTTCATCCGGAATGCTCCAGGTTCATTTCTGAGATTCTCATCAGGCA 1388
Qy 961 CACCAACAGCAGGACGAG 979
Db 1389 CACCAACAGCAGGACGAG 1407

RESULT 8
AB079066
LOCUS AB079066 4445 bp mRNA linear PRI 27-JUN-2002
DEFINITION Homo sapiens cerk mRNA for ceramide kinase, complete cds.
ACCESSION AB079066
VERSION AB079066.1 GI:21624339
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Suglura, M., Kono, K., Liu, H., Shimizugawa, T., Minekura, H.,
Spiegel, S. and Kohana, T.
Ceramide kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND
FUNCTIONAL CHARACTERIZATION
J. Biol. Chem. 277 (26), 23294-23300 (2002)
11956206
2 (bases 1 to 4445)
Sugira, M., Kono, K., Shimizugawa, T., Minekura, H., Spiegel, S. and
Kohama, T.
Direct Submission
Submitted (29-JAN-2002) Masako Sugira, Sankyo Co., Ltd.,
Pharmacology and Molecular Biology Research Laboratories; 2-58
Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan
(E-mail: msugira@shina.sankyo.co.jp, Tel: 81-3-3492-3131,
Fax: 81-3-5436-8565)
FEATURES
Location/Qualifiers
1..4445
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..4445
/gene="cerk"
124..1737
/gene="cerk"
/codon_start=1
/product="ceramide kinase"
/protein_id="BAC01154.1"
/db_xref="GI:21624340"
/translation="MGATGAAPLQSVLWVQRCVSLPAPALLRWRSPGPGAGA
PGADACSVPESEIITAVEETDVHGKQSGKWKMEKPYATVHCVKRARRHRWKAQV
TFWCPPEOLCHLQTLREMLEKLTSPKHLVFINPFGKGQKRIYERKVAPLFTL
ASITDIIVTHANQAKETLYEINIDKYDGLVCGDGMFSEVLHGLIGRTQASGVD
QNHPRVLPSSLRIGIIPAGSDVCYSTVGTSDAETSAHLIIVGDSLMDVSSVHH
NSTLRYVSLGTFYGDIIKSEKKRWGLARYDFSGKLTFLSHCYEGTVGFLPA
QHTVGSPRDRPCRAGCFVCRSQKQLEEQKALYGLEAAEDVEEWQVVGKFLAIN
ATNMSCACRRSPRGLSPAHLHGDSGLILIRKCSRFNFLRFLIRHTNQDDPDTFV
EYRVKKEQFTSKMEDESDLKSGGKRFGHCSSHPSCCTTSNSNSWNCDEVLHS
PAIEVRVHCQLVRLFARGIENPKPDSHS"

Query Match 92.2%; Score 903; DB 9; Length 4445;
Best Local Similarity 96.6%; Pred. No. 6.9e-229;
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;

Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTCGAGAAAGACAAAGCAAGCGGAT 60
Db 510 ACCAAGCATTTACTGGTATTATCAACCCGTTTCGAGAAAGACAAAGCAAGCGGAT 569
Qy 61 ATATGAAGAAAGTGGCACCACCTGTTTCACTTAGCTTCCATCACCCTGACATCATCGG 120
```

ORIGIN

```
RESULT 9
AR541900
LOCUS AR541900 4432 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 148 from patent US 6743619.
ACCESSION AR541900
VERSION AR541900.1 GI:53933980
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
```

```
Db 570 ATATGAAGAAAGTGGCACCACCTGTTTCACTTAGCTTCCATCACCCTGACATCATCG- 628
Qy 121 TAACAAATTCCTATGTTAACTATATGTAGAAGTAATTAATCTGAACATCTAATCAGGCCAAGGA 180
Db 629 -----TTACTGAAATGCTAATCAGGCCAAGGA 656
Qy 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCCGGCGGAGATGG 240
Db 657 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCCGGCGGAGATGG 716
Qy 241 TATGTTTCCAGGAGGTGTCACCGTCTGATTGGGAGGACGCGAGAGGAGCGCCGGGTCGA 300
Db 717 TATGTTTCCAGGAGGTGTCACCGTCTGATTGGGAGGACGCGAGAGGAGCGCCGGGTCGA 776
Qy 301 CCAGAACACACCCCGGGCTGTGTGGTCCCAGTAGCTCCGGATTCCGGAATGGAATCATTTCCCGC 360
Db 777 CCAGAACACACCCCGGGCTGTGTGGTCCCAGTAGCTCCGGATTGGAATCATTTCCCGC 836
Qy 361 AGGCTCAACGAGCTGCGTGTGTTACTCCACGTTGGGCACGAGCGACGCAAGAACTCCGC 420
Db 837 AGGCTCAACGAGCTGCGTGTGTTACTCCACGTTGGGCACGAGCGACGCAAGAACTCCGC 896
Qy 421 GCTGCATATCGTGTGGGACTCGCTGGCCATGGATGTGTCTCCTCAGTCCACCAACACAG 480
Db 897 GCTGCATATCGTGTGGGACTCGCTGGCCATGGATGTGTCTCCTCAGTCCACCAACACAG 956
Qy 481 CACACTCTTCGCTACTCCGTTCCCTGTGGGCTACGGCTTTTACGGGGACATCATCAA 540
Db 957 CACACTCTTCGCTACTCCGTTCCCTGTGGGCTACGGCTTTTACGGGGACATCATCAA 1016
Qy 541 GGACAGTGAAGAAACGGTGGTGGTCTTGCAGATACGACTTTTTCAGGTTTAAAGAC 600
Db 1017 GGACAGTGAAGAAACGGTGGTGGTCTTGCAGATACGACTTTTTCAGGTTTAAAGAC 1076
Qy 601 CTTCTCTTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCCTGTCACCAACACACGGT 660
Db 1077 CTTCTCTTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCCTGTCACCAACACACGGT 1136
Qy 661 GGGATCTCCAAGGGATAGGAAGCCCTCCCGGGCAGGATGCTTTGTTTCAGCGCAAGCAAA 720
Db 1137 GGGATCTCCAAGGGATAGGAAGCCCTCCCGGGCAGGATGCTTTGTTTCAGCGCAAGCAAA 1196
Qy 721 GCACAGCTGGAGGAGGACGAAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGT 780
Db 1197 GCACAGCTGGAGGAGGACGAAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGT 1256
Qy 781 GGAGGAGTGGCAAGTCTGTTGGGAAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTG 840
Db 1257 GGAGGAGTGGCAAGTCTGTTGGGAAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTG 1316
Qy 841 TGCTTGTCCCGGAGCCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTTC 900
Db 1317 TGCTTGTCCCGGAGCCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTTC 1376
Qy 901 TGACCTCATCTCATCCGAAAATGCTCAGGTTCAATTTCTGAGATTCTCATCAGGCA 960
Db 1377 TGACCTCATCTCATCCGAAAATGCTCAGGTTCAATTTCTGAGATTCTCATCAGGCA 1436
Qy 961 CACCAACAGCAGGACGAG 979
Db 1437 CACCAACAGCAGGACGAG 1455
```

| | | |
|---------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|
| REFERENCE | 1 | (bases 1 to 4432) |
| AUTHORS | Tang,Y.T., Zhou,P., Goodrich,R., Liu,C., Asundi,V., Ren,F., Zhang,J., Zhao,Q.A., Yang,Y., Xue,A.J., Wehrman,T., Wang,J.-R., Wang,D. and Drmanac,R.T. | |
| TITLE | Nucleic acids and polypeptides | |
| JOURNAL | Patent: US 6743619-A 148 01-JUN-2004; | |
| FEATURES | Location/Qualifiers | |
| source | 1..4432 | |
| ORIGIN | /organism="unknown"
/mol_type="genomic DNA" | |
| Query Match | 92.1%; Score 901.4; DB 6; Length 4432; | |
| Best Local Similarity | 96.5%; Pred. No. 1.8e-228; | |
| Matches 945; Conservative | 0; Mismatches 1; Indels 33; Gaps 1; | |
| Qy | 1 ACCAAAGCATTTACTGTGTTATTTATCAACCCGCTTTGGAGGAAAAGGACAAGCGCAAGCGGAT 60 | |
| Db | 1590 ACCAAAGCATTTACTGTGTTATTTATCAACCCGCTTTGGAGGAAAAGGACAAGCGCAAGCGGAT 1649 | |
| Qy | 61 ATATGAAGAAAAGTGGCAACACTGTTCACTTAGCCTCCATCACCACTGACATCATCGG 120 | |
| Db | 1650 ATATGAAGAAAAGTGGCAACACTGTTCACTTAGCCTCCATCACCACTGACATCATCG- 1708 | |
| Qy | 121 TAACAAAATCTATGTTAACTATGTAGAAGTAATTACTGAAATGCTTAATCAGGCCAAGGA 180 | |
| Db | 1709 -----TTACTGAACATGCTTAATCAGGCCAAGGA 1736 | |
| Qy | 181 GACTCTGTATGAGATTAACTAGACAATACGACGGCATCTGTGTCTGGCGGAGATGG 240 | |
| Db | 1737 GACTCTGTATGAGATTAACTAGACAATACGACGGCATCTGTGTCTGGCGGAGATGG 1796 | |
| Qy | 241 TATGTTTCAGGAGGTGCTGCACGCTGTGATTGGGAGGACGCAGAGGAGCGCCGGGGTCGA 300 | |
| Db | 1797 TATGTTTCAGGAGGTGCTGCACGCTGTGATTGGGAGGACGCAGAGGAGCGCCGGGGTCGA 1856 | |
| Qy | 301 CCAGAACCAACCCCGGGCTGTGTGGTCCCAAGTAGCCTCGGATTGGAATCATTTCCCGC 360 | |
| Db | 1857 CCAGAACCAACCCCGGGCTGTGTGGTCCCAAGTAGCCTCGGATTGGAATCATTTCCCGC 1916 | |
| Qy | 361 AGGGTCAACGGACTCGTGTTACTCCACGGTGGGACACGACGACCGAGAACCTCGGC 420 | |
| Db | 1917 AGGGTCAACGGACTCGTGTTACTCCACGGTGGGACACGACGACCGAGAACCTCGGC 1976 | |
| Qy | 421 GCTGCATATCGTTGTGGGACTCGCTGGCCATGGATGTCTCAGTCAACACCAACAG 480 | |
| Db | 1977 GCTGCATATCGTTGTGGGACTCGCTGGCCATGGATGTCTCAGTCAACACCAACAG 2036 | |
| Qy | 481 CACACTCTTCGCTACTCCGTTCCCTGTGGGCTACGGGTTCTACGGGGACATCATCAA 540 | |
| Db | 2037 CACACTCTTCGCTACTCCGTTCCCTGTGGGCTACGGGTTCTACGGGGACATCATCAA 2096 | |
| Qy | 541 GGACAGTGAGAAGAAAACGGTGGTGGGTCTTGCCAGATACGACTTTTTCAGTTTAAAGAC 600 | |
| Db | 2097 GGACAGTGAGAAGAAAACGGTGGTGGGTCTTGCCAGATACGACTTTTTCAGTTTAAAGAC 2156 | |
| Qy | 601 CTTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACACGGT 660 | |
| Db | 2157 CTTTCTCTCCACCACTGCTATGAGGACAGTGTCTTCTCTCTCCCTGCACAAACACACGGT 2216 | |
| Qy | 661 GGGATCTCCAAGGATAGGAGCCCTCGCGGACAGGATGCTTTGTTTTCAGGCGAAGCAA 720 | |
| Db | 2217 GGGATCTCCAAGGATAGGAGCCCTCGCGGACAGGATGCTTTGTTTTCAGGCGAAGCAA 2276 | |
| Qy | 721 GCAGCAGCTGGAGGAGGACGAGAAGAAGCACTGTATGTTTGGAAAGCTGCGGAGGACGT 780 | |
| Db | 2277 GCAGCAGCTGGAGGAGGACGAGAAGAAGCACTGTATGTTTGGAAAGCTGCGGAGGACGT 2336 | |
| Qy | 781 GGAGGAGTGCAAGTCTGTGTGGAGTTTCTGGCCATCAATGCCCAACAACATGTCCTG 840 | |
| Db | 2337 GGAGGAGTGCAAGTCTGTGTGGAGTTTCTGGCCATCAATGCCCAACAACATGTCCTG 2396 | |
| Qy | 841 TGCTTTGTGCGCGGAGCCCGAGGGGCTCTCTCCCGGCTGCCACATTTGGAGACGGGCTCTTC 900 | |

Db 2397 TGCTTGTGCGCGAGCCCGAGGGGCGCTCTCCCGGCTGCCACTTGGAGACGGGTCTTC 2456
 Qy 901 TGACCTCATCTCATCCGGAATAGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960
 Db 2457 TGACCTCATCTCATCCGGAATAGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 2516
 Qy 961 CACCAACGAGGAGACCAG 979
 Db 2517 CACCAACGAGGAGACCAG 2535
 RESULT 10
 BD102675
 LOCUS Ceramide kinase and DNA thereof. 4463 bp DNA linear PAT 27-AUG-2002
 DEFINITION
 ACCESSION BD102675
 VERSION BD102675.1 GI:22648249
 KEYWORDS WO 0196575-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4463)
 AUTHORS Sugiura,M., Kono,K. and Kohama,T.
 TITLE Ceramide kinase and DNA thereof
 JOURNAL Patent: WO 0196575-A 1 20-DEC-2001;
 SANKYO CO LTD,MASAKO SUGIURA,KEITA KONO,TAKAFUMI KOHAMA
 COMMENT OS Homo sapiens (human)
 PN WO 0196575-A/1
 PD 20-DEC-2001
 PF 11-JUN-2001 WO 2001JP004889
 PR 14-JUN-2000 JP 00P 178039
 PI MASAKO SUGIURA,KEITA KONO,TAKAFUMI KOHAMA
 PC C12N15/54,C12N9/12,C12N1/21,C12Q1/48,C07K16/40,A61K31/7125,PC
 A61K48/00,
 PC A61P3/04,A61P3/10,A61P7/00,A61P25/00,A61P29/00,A61P31/18,PC
 A61P35/00
 CC Ceramide kinase and DNA thereof
 FH Key Location/Qualifiers
 FT CDS (124)..(1734).
 FEATURES
 source 1..4463
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 91.7%; Score 898.2; DB 6; Length 4463;
 Best Local Similarity 96.3%; Pred. No. 1.3e-227;
 Matches 943; Conservative 0; Mismatches 3; Indels 33; Gaps 1;
 Qy 1 ACCAAACATTTTACTGCTATTATTATCAACCGTTTGGAGGAAAGCAAGCAAGCGCGAT 60
 Db 510 ACCAAACATTTTACTGCTATTATTATCAACCGTTTGGAGGAAAGCAAGCAAGCGCGAT 569
 Qy 61 ATATGAAGAAAAGTGGCACCACTGTTTCACCTTAGCCTCCATCACCACTGACATCATCGG 120
 Db 570 ATATGAAGAAAAGTGGCACCACTGTTTCACCTTAGCCTCCATCACCACTGACATCATCG- 628
 Qy 121 TAACAAATTTCTATGTTAACTATGTAGAGTAATTTACTTGAACATGCTAATCAGGCCAAGGA 180
 Db 629 -----TTACTGAACATGCTAATCAGGCCAAGGA 656
 Qy 181 GACTCTCTATGAGATTAACTAGACAAATAACGCGCATCGTCTGTGTCGCGGAGATGG 240
 Db 657 GACTCTCTATGAGATTAACTAGACAAATAACGCGCATCGTCTGTGTCGCGGAGATGG 716
 Qy 241 TATGTTTCAGGAGTGTGTCACCGTCTGATTGGAGCACCGAGGAGCCCGGGGTGCA 300
 Db 717 TATGTTTCAGGAGTGTGTCACCGTCTGATTGGAGGACCGAGGAGCCCGGGGTGCA 776
 Qy 301 CCAGAACACACCCCGGGCTGTGTGTGTTCCCAAGTAGCCTCCGGAATTGGAATCATTTCCCGC 360

|||||
777 CAGAACACCCCGGCTGTCTGGTCCCGAGTAGCTCCGGATTGGAATCATTTCCGCG 836
361 AGGCTCAACGGACTCGTGTGTACTCCACCGTGGGACACGAGCGAGCAAACTCGGC 420
837 AGGTCCAACGGACTCGTGTGTACTCCACCGTGGGACACGAGCGAGCAAACTCGGC 896
421 GCTGATATCTGTGTGGGACTCGTGGCCATGATGTCTCAGTCCACCAACAG 480
897 GCTGATATCTGTGTGGGACTCGTGGCCATGATGTCTCAGTCCACCAACAG 956
481 CACACTCTCTCCCTACTCTCGTCTCCCTGTGGGCTACGCTTCTACGGGACATCATCA 540
957 CACACTCTCTCCCTACTCTCGTCTCCCTGTGGGCTACGCTTCTACGGGACATCATCA 1016
541 GGACAGTGAGAAACCGTGTGGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGAC 600
1017 GGACAGTGAGAAACCGTGTGGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGAC 1076
601 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCGTCACAAACACAGGT 660
1077 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCGTCACAAACACAGGT 1136
661 GGGATCTCAAGGGATAGGAAGCCCTGCGGGCAGATGCTTTGTTGACGGCAAGCAA 720
1137 GGGATCTCAAGGGATAGGAAGCCCTGCGGGCAGATGCTTTGTTGACGGCAAGCAA 1196
721 GCAGCAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAGCTGCGGAGGACGT 780
1197 GCAGCAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAGCTGCGGAGGACGT 1256
781 GGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGCCATCAATGCCACAAACATGTCCTG 840
1257 GGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGCCATCAATGCCACAAACATGTCCTG 1316
841 TGCTTGTGCGGGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTC 900
1317 TGCTTGTGCGGGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTC 1376
901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCATCAGGCA 960
1377 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCATCAGGCA 1436
961 CACCAACAGCAGGACCAG 979
1437 CACCAACAGCAGGACCAG 1455

RESULT 11
CQ730476 1459 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 16410 from Patent WO02068579.
DEFINITION CQ730476
ACCESSION CQ730476
VERSION CQ730476.1 GI:42304409
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
AUTHORS Kite, such as nucleic acid arrays, comprising a majority of
TITLE humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 16410 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source
1..1459
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 84.5%; Score 826.8; DB 6; Length 1459;
Best Local Similarity 99.8%; Pred. No. 1.1e-208;
Matches 828; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 150 TAAATTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAT 209
DB |||||||
DB 348 TGATTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAT 407
QY 210 ACGACGGCATCGTCTGTGTGCGCGAGATGTATTTCAGCGAGGTGCTGACGGTCTGA 269
DB |||||||
DB 408 ACGACGGCATCGTCTGTGTGCGCGAGATGTATTTCAGCGAGGTGCTGACGGTCTGA 467
QY 270 TTGGGAGGACGACAGAGCGCGCGGCTCGACAGAACCCCGGGCTGTGCTGGTCC 329
DB |||||||
DB 468 TTGGGAGGACGACAGAGCGCGCGGCTCGACAGAACCCCGGGCTGTGCTGGTCC 527
QY 330 CCAGTAGCCTCCGATTGGGAATCATTTCCGCGAGGCTCAACGAGCTGCGTGTGTTACTCCA 389
DB |||||||
DB 528 CCAGTAGCCTCCGATTGGGAATCATTTCCGCGAGGCTCAACGAGCTGCGTGTGTTACTCCA 587
QY 390 CCGTGGGACACGACGCGACGAGAAACCTCGGCGCTGCATATCGTTTGGGAGCTCGCTGG 449
DB |||||||
DB 588 CCGTGGGACACGACGCGACGAGAAACCTCGGCGCTGCATATCGTTTGGGAGCTCGCTGG 647
QY 450 CCATGGATGTCTCTCAGTCCACCAACAGCACTCTTCCGCTACTCCGTTCCCTGCC 509
DB |||||||
DB 648 CCATGGATGTCTCTCAGTCCACCAACAGCACTCTTCCGCTACTCCGTTCCCTGCC 707
QY 510 TGGGCTACGGCTTCTAGCGGACATCATCAAGGACAGTGAAGAAACGCTGGTGGGTGTC 569
DB |||||||
DB 708 TGGGCTACGGCTTCTAGCGGACATCATCAAGGACAGTGAAGAAACGCTGGTGGGTGTC 767
QY 570 TTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCCACTGCTATGAAGGGA 629
DB |||||||
DB 768 TTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCCACTGCTATGAAGGGA 827
QY 630 CAGTGTCTCTCTCCCTGCACAAACAGCGTGGGATCTCCAGGATAGGAAGCCCTGCC 689
DB |||||||
DB 828 CAGTGTCTCTCTCCCTGCACAAACAGCGTGGGATCTCCAGGATAGGAAGCCCTGCC 887
QY 690 GGGCAGGATGCTTTGTTTGGAGGCAACAGCAGCAGCTGGAGAGGAGCAGAGAAAG 749
DB |||||||
DB 888 GGGCAGGATGCTTTGTTTGGAGGCAACAGCAGCAGCTGGAGAGGAGCAGAGAAAG 947
QY 750 CACTGTATGTTTGGAAAGCTCGGAGGACGTTGGAGGAGTGGCAAGTCTGTGGGAAGT 809
DB |||||||
DB 948 CACTGTATGTTTGGAAAGCTCGGAGGACGTTGGAGGAGTGGCAAGTCTGTGGGAAGT 1007
QY 810 TTCTGGCCATCAATGCCACAAACATGTCTGTGCTTGTGCGCGAGCCCCAGGGCCCTCT 869
DB |||||||
DB 1008 TTCTGGCCATCAATGCCACAAACATGTCTGTGCTTGTGCGCGAGCCCCAGGGCCCTCT 1067
QY 870 CCCCAGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAATGCTCCA 929
DB |||||||
DB 1068 CCCCAGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAATGCTCCA 1127
QY 930 GGTTCAATTTTCTGAGATTTCATCAGGACACCAACACAGCAGGACCAG 979
DB |||||||
DB 1128 GGTTCAATTTTCTGAGATTTCATCAGGACACCAACACAGCAGGACCAG 1177

RESULT 12
BC067255 1772 bp mRNA linear PRI 06-APR-2004
LOCUS Homo sapiens ceramide kinase, mRNA (cdna clone IMAGE:6185601),
DEFINITION complete cds.
ACCESSION BC067255
VERSION BC067255.1 GI:45595582
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

| | |
|----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| REFERENCE
AUTHORS | 1 (bases 1 to 1772)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heideh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavanti,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEvan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shevchenko,Y.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Green,E.D.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalilus,D.E.,
Schnerth,A., Schein,J.E., Jones,S.J. and Marra,M.A. |
| TITLE | Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) |
| REFERENCE
AUTHORS | 2 (bases 1 to 1772)
Strausberg,R. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (12-MAR-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA |
| REMARK
COMMENT | NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palquist, Anca Petrescu, Anna Liisa Prabh,
Parvaneh Saeeedi, JR Santos, Angeliue Schnerch, Ursula Skalska,
Duane Smalilus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
Schein, Asim Siddiqui, Rob Holt, Marco Marra. |
| FEATURES
source | Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 141 Row: d Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21703365
This clone has the following problem: The cds is short compared to
the longest cds in the locus.
Location/Qualifiers
1..1772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6185601"
/tissue_type="Peripheral Nervous System, dorsal root
ganglion"
/clone_lib="Lupski_dorsal_root_ganglion"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1..1772
/gene="CERK"
/note="synonyms: KIAA1646, hCERK, dA59H18.3,
DKFZp434E0211, FLJ21430, FLJ23239, dA59H18.2, LK4"
/db_xref="LocusID:64781"
579..1598
/gene="CERK"
/codon_start=1 |
| gene | ORIGIN
Query Match 78.1%; Score 764.8; DB 9; Length 1772;
Best Local Similarity 99.1%; Pred. No. 3.7e-192;
Matches 769; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 204 ACAATACGAGCGATCGTCTGTGTCGGCGAGATGATGTTTCAGCGAGGTGTCGACG 263
Db 541 ACTGACATCATCGCATCGTCTGTGTCGGCGAGATGATGTTTCAGCGAGGTGTCGACG 600
Qy 264 GTCTGATTGGAGACGACGAGGAGCGCGGGTTCGACCAAGACCAACCCCGGCTGTCG 323
Db 601 GTCTGATTGGAGACGACGAGGAGCGCGGGTTCGACCAAGACCAACCCCGGCTGTCG 660
Qy 324 TGGTCCCAGTAGCTCCGGATTGGAATCATTCCTCCGAGGGTCAACCGACTGCGTGTGTT 383
Db 661 TGGTCCCAGTAGCTCCGGATTGGAATCATTCCTCCGAGGGTCAACCGACTGCGTGTGTT 720
Qy 384 ACTCCACGCTGGGACGACGAGCGAGAAACCTTCGGCGCTGCATATCGTTGTTGGGACT 443
Db 721 ACTCCACGCTGGGACGACGAGCGAGAAACCTTCGGCGCTGCATATCGTTGTTGGGACT 780
Qy 444 CGCTGGCGATGGATGTCCTCAGTCCACCAACAGCACACTCTTCGCTACTCCGTGT 503
Db 781 CGCTGGCGATGGATGTCCTCAGTCCACCAACAGCACACTCTTCGCTACTCCGTGT 840
Qy 504 CCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGAAGAAACGGTGGT 563
Db 841 CCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGAAGAAACGGTGGT 900
Qy 564 TGGGCTTTCGACAGATACGACTTTTCAGTTTAAAGACCTTCCTCCACACATGCTATG 623
Db 901 TGGGCTTTCGACAGATACGACTTTTCAGTTTAAAGACCTTCCTCCACACATGCTATG 960
Qy 624 AAGGACAGTGTCTCTCCCTGACACACACAGTGGGATCTTCAAGGGATAGGAAGC 683
Db 961 AAGGACAGTGTCTCTCCCTGACACACACAGTGGGATCTTCAAGGGATAGGAAGC 1020
Qy 684 CCTCCCGGCGAGGATGCTTTGTTTGAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGA 743
Db 1021 CCTCCCGGCGAGGATGCTTTGTTTGAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGA 1080
Qy 744 AGAAGACATGTATGTTTGAAGCTCGGAGGACGTGGAGGAGTGCAGTCTCTGTG 803
Db 1081 AGAAGACATGTATGTTTGAAGCTCGGAGGACGTGGAGGAGTGCAGTCTCTGTG 1140
Qy 804 GGAAGTTTCTGGCCATCAATGCCAACATGTCTCTGTTGTCGGCGAGGCCACAGG 863
Db 1141 GGAAGTTTCTGGCCATCAATGCCAACATGTCTCTGTTGTCGGCGAGGCCACAGG 1200
Qy 864 GCCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAT 923
Db 1201 GCCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAT 1260
Qy 924 GCTCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAACACAGCAGGACACAG 979
Db 1261 GCTCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAACACAGCAGGACACAG 1316 |
| CDS | RESULT 13
AKI29416
LOCUS AKI29416 3661 bp mRNA linear ROD 21-NOV-2003 |

```

DEFINITION Mus musculus mRNA for mKIAA1646 protein.
ACCESSION AK129416
VERSION AK129416.1 GI:37360495
KEYWORDS FLI CDNA.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Koseki,H., Hiraoka,S.,
Saga,Y., Nagase,T., Ohara,O. and Koga,H.
TITLE Prediction of the coding sequences of mouse homologues of KIAA
gene: III. the complete nucleotide sequences of 500 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
JOURNAL DNA Res. 10 (4), 167-180 (2003)
MEDLINE 22977043
PUBMED 14621295
REFERENCE
2 (bases 1 to 3661)
AUTHORS Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2003) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing; Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
FEATURES
Location/Qualifiers
source
1..3661
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mph01246"
/tissue_type="embryonic tail"
/note="vector:modified pBC SK+"
1..3661
/gene="mKIAA1646"
<1..1231
/gene="mKIAA1646"
/note="CDS is predicted by in silico analysis. Start codon
is not identified.
Start codon is not identified."
/codon_start=2
/evidence="not experimental"
/product="mKIAA1646 protein"
/protein_id="BAC98226.1"
/db_xref="GI:37360496"
/translation="FCFAASRPKHLVFINPFGKGQKRIYKTVAPLFTLASITTE
IIITEHANOAKETIYEINTSDYGVICVGGDMFSEVLHVGIRTOQSAGIDPNHRA
VLVPSLTIGIIPAGSTDCVSVTGTNDASTSALHIIIGDSLAIIDSVSVHYHNTLLR
YVSLGLGYFGDLIKDSEKKRWMLVRYDFSLGKTLFSLHOVYETLSFLPAQHTVGS
PRDNKPCPAGCFVCSKOOLEEKKALYGLNDAEEVEEWQVTCGFLANATNMSC
ACPRSPGSLFPAHUGSSDILLIRKSRNFRLRHLRHNTQEDQDFTFVFEYVRK
VHFQVRLFLFARGIEES"
ORIGIN
Query Match 65.3%; Score 639; DB 10; Length 3661;
Best Local Similarity 79.8%; Pred. No. 1.2e-158;
Matches 781; Conservative 0; Mismatches 165; Indels 33; Gaps 1;
QY 1 ACCAAGACTTTACTGGTATTATCAACCCGGTTTGGAGGAAAGGACAAAGCGGAT 60
Db 22 ACCGAGGACCTTGCTGGTATTATCAACCCCTTCGAGGGAAGGTGAGGCGACGGAT 81
QY 61 ATATGAAGAAAAGTGGCACCCTGTTCACCTTAGCTCCATCCACTGACATCATCGG 120
Db 82 CTATGAAAAACAGTGGCGCTCTGTTTACCTTGGCTTCCATCACTACGGAGATCATC- 139
QY 121 TAACAAATCTATGTAACTATGTAGAAGTAACTTACTGAACATGTATATCAGGCCAAGA 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```
QY 359 GCAGGTCACGGACTCGTGTGTTACTCCACCGTGGGCACGACGACGAGAACCTCG 418
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 128 GCAGGTCACGGACTCGTGTGTTACTCCACCGTGGGCACGACGACGAGAACCTCG 187
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 419 GCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAAC 478
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 188 GCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAAC 247
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 479 AGCACACTCCTTCGCTACTCCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATC 538
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 248 AGCACACTCCTTCGCTACTCCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATC 307
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 539 AAGGACAGTGAGAAACGGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAG 598
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 308 AAGGACAGTGAGAAACGGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAG 367
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 599 ACCTTCTCTCCACCACTGTATGAAGGGACAGTGTCTTCTCCTCCCTGCACAAACACG 658
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 368 ACCTTCTCTCCACCACTGTATGAAGGGACAGTGTCTTCTCCTCCCTGCACAAACACG 427
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 659 GTGGGATCTCCAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTG 704
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 428 GTGGGATCTCCAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTG 473
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Search completed: September 5, 2005, 23:41:36
Job time : 4461.35 secs

2.

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 15:04:18 ; Search time 564.067 Seconds
(without alignments)
10274.357 Million cell updates/sec

Title: US-10-631-958-1
Perfect score: 979
Sequence: 1 accaaagcattactgtat.....acaccaaccagcaggaccag 979

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_16Dec04:*
- 1: Geneseqn1980s:*
 - 2: Geneseqn1990s:*
 - 3: Geneseqn2000s:*
 - 4: Geneseqn2001as:*
 - 5: Geneseqn2001bs:*
 - 6: Geneseqn2002as:*
 - 7: Geneseqn2002bs:*
 - 8: Geneseqn2003as:*
 - 9: Geneseqn2003bs:*
 - 10: Geneseqn2003cs:*
 - 11: Geneseqn2003ds:*
 - 12: Geneseqn2004as:*
 - 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 979 | 100.0 | 979 | 6 | ABL40822 Human sph |
| 2 | 979 | 100.0 | 1840 | 5 | Aad14426 Human sph |
| 3 | 903 | 92.2 | 1740 | 8 | Ada05679 Human NOV |
| 4 | 903 | 92.2 | 3975 | 8 | Acc70838 Human Sph |
| 5 | 903 | 92.2 | 4231 | 3 | Aaa50510 Human sph |
| 6 | 903 | 92.2 | 4413 | 6 | Ab140828 Human sph |
| 7 | 903 | 92.2 | 4429 | 12 | Aqj96598 Human lip |
| 8 | 903 | 92.2 | 4445 | 13 | Adp55247 Human PRO |
| 9 | 901.4 | 92.1 | 4432 | 8 | Abx70921 Novel hum |
| 10 | 898.2 | 91.7 | 4463 | 6 | Aba96945 Human cer |
| 11 | 878.4 | 89.7 | 1740 | 12 | Adn62844 Human NOV |
| 12 | 622.2 | 63.6 | 2241 | 5 | Aas77728 DNA encod |
| 13 | 618 | 63.1 | 4702 | 13 | Adsl10370 Human the |
| 14 | 598.8 | 61.2 | 753 | 3 | Aac76031 Human ORF |
| 15 | 543 | 55.5 | 1570 | 5 | Aas77731 DNA encod |
| 16 | 543 | 55.5 | 2186 | 5 | Aas77730 DNA encod |
| 17 | 464.4 | 47.4 | 474 | 6 | ABL40823 Human sph |
| 18 | 401.8 | 41.0 | 426 | 5 | Aas77727 DNA encod |
| 19 | 362.8 | 37.1 | 411 | 5 | Aas77729 DNA encod |
| 20 | 302.2 | 30.9 | 329 | 6 | ABL40824 Human sph |

| | | | | | | |
|----|-------|------|-------|----|----------|---------------------|
| 21 | 241 | 24.6 | 547 | 3 | AAC76592 | Aac76592 Human ORF |
| 22 | 209.8 | 21.4 | 15181 | 4 | AAK65588 | Aak65588 Human imm |
| 23 | 209.8 | 21.4 | 15185 | 4 | AAK65589 | Aak65589 Human imm |
| 24 | 200.6 | 20.5 | 564 | 8 | ABT23453 | Abt23453 Immune-re |
| 25 | 167 | 17.1 | 167 | 6 | ABL40825 | Ab140825 Human sph |
| 26 | 153 | 15.6 | 153 | 6 | ABL40826 | Ab140826 Human sph |
| 27 | 140 | 14.3 | 817 | 6 | ABQ99499 | Abq99499 Human cod |
| 28 | 140 | 14.3 | 817 | 13 | ADS11669 | Adsl1669 Human the |
| 29 | 136.8 | 14.0 | 522 | 5 | AAD14427 | Aad14427 Partial r |
| 30 | 89 | 9.1 | 550 | 6 | ABL40827 | Ab140827 Human sph |
| 31 | 77 | 7.9 | 2173 | 4 | ABL25705 | Ab125705 Drosophil |
| 32 | 77 | 7.9 | 10337 | 4 | ABL25704 | Ab125704 Drosophil |
| 33 | 69.8 | 7.1 | 498 | 6 | ABK33019 | Abk33019 DNA encod |
| 34 | 57.4 | 5.9 | 2000 | 8 | ADA71938 | Ada71938 Rice gene |
| 35 | 52.2 | 5.3 | 1774 | 12 | ADQ88890 | Adq88890 Novel hum |
| 36 | 51 | 5.2 | 763 | 8 | ABT23456 | Abt23456 Immune-re |
| 37 | 49.4 | 5.0 | 2000 | 8 | ADA71938 | Ada71938 Rice gene |
| 38 | 43.4 | 4.4 | 1050 | 11 | ABD02665 | Abd02665 Pseudomon |
| 39 | 43.4 | 4.4 | 1308 | 11 | ABD02575 | Abd02575 Pseudomon |
| 40 | 41.8 | 4.3 | 1167 | 4 | AAF81355 | Aaf81355 Quorum se |
| 41 | 41.8 | 4.3 | 1167 | 13 | ADS14656 | Adel14656 Pseudomon |
| 42 | 40.4 | 4.1 | 1857 | 6 | ABK90199 | Abk90199 cDNA enco |
| 43 | 40.4 | 4.1 | 2380 | 5 | AAS14817 | Aas14817 Human cDN |
| 44 | 40.4 | 4.1 | 2380 | 6 | ABL59533 | Ab159533 Human sph |
| 45 | 40.4 | 4.1 | 2380 | 10 | ADE85298 | Ade85298 Farnesyl |

ALIGNMENTS

| | | | | | |
|----------|---------------------------------------------------------------------------|--|--|--|--|
| RESULT 1 | | | | | |
| ABL40822 | | | | | |
| ID | ABL40822 standard; cDNA; 979 BP. | | | | |
| XX | XX | | | | |
| AC | ABL40822; | | | | |
| XX | XX | | | | |
| DT | 03-JUL-2002 (first entry) | | | | |
| XX | XX | | | | |
| DE | Human sphingosine kinase-like protein encoding cDNA. | | | | |
| XX | XX | | | | |
| KW | Human sphingosine kinase-like protein; intracellular signalling; gene; | | | | |
| KW | cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma; | | | | |
| KW | autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss. | | | | |
| XX | XX | | | | |
| OS | Homo sapiens. | | | | |
| XX | XX | | | | |
| PH | Key | | | | |
| CDS | Location/Qualifiers | | | | |
| FT | 2..789 | | | | |
| FT | /*tag= a | | | | |
| FT | /product= "sphingosine kinase-like protein" | | | | |
| FT | /note= "start and stop codons are not indicated" | | | | |
| XX | XX | | | | |
| PN | WO200228906-A2. | | | | |
| XX | XX | | | | |
| PD | 11-APR-2002. | | | | |
| XX | XX | | | | |
| PF | 05-OCT-2001; 2001WO-EP011516. | | | | |
| XX | XX | | | | |
| PR | 06-OCT-2000; 2000US-02380005P. | | | | |
| PR | 23-AUG-2001; 2001US-0314113P. | | | | |
| XX | XX | | | | |
| PA | (FARB) BAYER AG. | | | | |
| XX | XX | | | | |
| PI | Kossida S, Encinas J; | | | | |
| XX | XX | | | | |
| DR | WPI; 2002-340094/37. | | | | |
| DR | P-PSDB; ABB07854. | | | | |
| XX | XX | | | | |
| PT | New reagent for modulating the activity of sphingosine kinase-like | | | | |
| PT | protein polypeptide or polynucleotide and treating cancer, asthma, | | | | |
| PT | allergy, an autoimmune disease, or a central or peripheral nervous system | | | | |
| PT | disorder. | | | | |
| XX | XX | | | | |

PS Claim 1; Fig 1; 120pp; English.

XX The invention relates to a human sphingosine kinase-like protein. The

CC polypeptide can be expressed by standard recombinant methodology. The

CC sphingosine kinase-like protein and gene can be used to regulate

CC intracellular signalling and consequently cell proliferation and

CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.

CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and

CC peripheral nervous system disorders (e.g. Parkinson's disease). The

CC present sequence represents the human sphingosine kinase-like protein

CC encoding cDNA

XX

SQ Sequence 979 BP; 237 A; 259 C; 271 G; 212 T; 0 U; 0 Other;

Query Match 100.08; Score 979; DB 6; Length 979;

Best Local Similarity 100.08; Pred. No. 2.3e-281;

Matches 979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCATTTACTGTTATTATCAACCCGTTTGGAGGAAAGGACAAGGCAAGCGGAT 60

DB 1 ACCAAGCATTTACTGTTATTATCAACCCGTTTGGAGGAAAGGACAAGGCAAGCGGAT 60

QY 61 ATATGAAAGAAAGTGGACCACTGTTCCACTTAGCCTTCCATCCACTGACATCATCGG 120

DB 61 ATATGAAAGAAAGTGGACCACTGTTCCACTTAGCCTTCCATCCACTGACATCATCGG 120

QY 121 TAACAAATTCATGTTAACTATGTAGAGTAATTAAGCATGCTTAATCAGGCCAAGGA 180

DB 121 TAACAAATTCATGTTAACTATGTAGAGTAATTAAGCATGCTTAATCAGGCCAAGGA 180

QY 181 GACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTGTGTGCGCGGAGATGG 240

DB 181 GACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTGTGTGCGCGGAGATGG 240

QY 241 TATGTTACGAGGTGCTGCACGGTCTGATTGGGAGGACGAGAGCGCGGGTCTGA 300

DB 241 TATGTTACGAGGTGCTGCACGGTCTGATTGGGAGGACGAGAGCGCGGGTCTGA 300

QY 301 CCAGAACACCCCGGGTGTGCTGGTCCCGTAGTACGCTCCGGATTGGAATCATTCGCGC 360

DB 301 CCAGAACACCCCGGGTGTGCTGGTCCCGTAGTACGCTCCGGATTGGAATCATTCGCGC 360

QY 361 AGGGTCAACGGAAGTGGTGTGTTACTCCACGTTGGGACACGACGAGAAACCTCGGC 420

DB 361 AGGGTCAACGGAAGTGGTGTGTTACTCCACGTTGGGACACGACGAGAAACCTCGGC 420

QY 421 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAG 480

DB 421 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAG 480

QY 481 CACACTCTTCGCTACTCCGTTCCCTGCTGGCTACGGCTTCTACGGGACATCATCA 540

DB 481 CACACTCTTCGCTACTCCGTTCCCTGCTGGCTACGGCTTCTACGGGACATCATCA 540

QY 541 GGACAGTGAGAGAAACGGTGGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 600

DB 541 GGACAGTGAGAGAAACGGTGGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 600

QY 601 CTTCTCTCCACCACTGCTATGAAAGGACAGTGTCTTCTCTCCCTGACAAACACCGGT 660

DB 601 CTTCTCTCCACCACTGCTATGAAAGGACAGTGTCTTCTCTCCCTGACAAACACCGGT 660

QY 661 GGGATCTCCAGGGATAGAGCCCTGCGGGCAGGATGCTTTGTTTCAGCGCAAGGAA 720

DB 661 GGGATCTCCAGGGATAGAGCCCTGCGGGCAGGATGCTTTGTTTCAGCGCAAGGAA 720

QY 721 GCAGCAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAGCTGCGGAGGACGT 780

DB 721 GCAGCAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAGCTGCGGAGGACGT 780

QY 781 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAATGTCCTG 840

DB 781 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAATGTCCTG 840

QY 841 TGCTTGTGCGGAGCCCGAGGGCCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 900

DB 841 TGCTTGTGCGGAGCCCGAGGGCCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 900

QY 901 TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960

DB 901 TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960

QY 961 CACCAACCAGCAGACCAG 979

DB 961 CACCAACCAGCAGACCAG 979

RESULT 2

AAD14426

ID AAD14426 standard; cDNA; 1840 BP.

XX

XX AAD14426;

AC AC

DT 01-NOV-2001 (first entry)

XX

XX Human sphingosine kinase (SphK) cDNA #2.

DE

XX

XX Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy;

KW antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;

KW cytosolic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;

KW leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss.

XX

OS Homo sapiens.

Key Location/Qualifiers

FF 109..1524

FT /*tag= a

FT /product= "Human sphingosine kinase (SphK) protein #2"

FT

FN WO200160990-A2.

XX

XX 23-AUG-2001.

XX

XX 14-FEB-2001; 2001WO-US004789.

XX

XX 14-FEB-2000; 2000US-0182360P.

PR 22-MAR-2000; 2000US-0191261P.

XX

XX (CURA-) CURAGEN CORP.

PA (GETH) GENENTECH INC.

XX

XX Rastelli L;

XX

XX WPI; 2001-514770/56.

DR P-PSDB; AAE07884.

XX

XX An isolated sphingosine kinase polypeptide useful for treating a SphK-

FT associated disorder especially cancer, restenosis or ischemia in a human.

XX

XX Claim 8; Page 94-95; 107pp; English.

XX

XX The present invention relates to sphingosine kinase (SphK) polypeptides

CC and nucleic acids encoding them. SphK is useful for treating a SphK-

CC associated disorder especially cancers such as leukaemia, lymphoma,

CC ovarian, breast, lung, colon, testicular, stomach and skin,

CC atherosclerosis, restenosis or ischaemia and cell proliferative disease

CC or disorder associated with vascular diseases. SphK gene is used in gene

CC therapy and antisense-therapy. Sphingolipids serving as signalling

CC molecules, have recently emerged as regulators of cell growth,

CC differentiation, diverse cell phenotypes and cell death. Activation of

CC SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human

CC endothelial cells. The present sequence is human sphingosine kinase

CC (SphK) cDNA

XX

SQ Sequence 1840 BP; 459 A; 468 C; 503 G; 410 T; 0 U; 0 Other;

| | | |
|---------------------------------------------------------------|---------------------------------------------------------------------------|--|
| PA | (CHBI-) CHEM BIOLOGY INST. | |
| XX | Igarashi Y, Kihara A; | |
| XX | WPI; 2003-354917/33. | |
| DR | P-PSDB; ABR56301. | |
| XX | Platelet derived polypeptides with sphingosine kinase activity for | |
| PT | treatment of sphingosine related disorders. | |
| XX | Claim 4; Page 30-31; 39pp; Japanese. | |
| XX | The present sequence is the coding sequence for human sphingosine kinase | |
| CC | 4 (SPHK4). The kinase can be used for the diagnosis and treatment of | |
| CC | sphingosine related disorders. The kinase can also be potentially used | |
| CC | for controlling toxicity of platelet transfusion and as a platelet | |
| CC | stabiliser | |
| XX | Sequence 3975 BP; 943 A; 978 C; 1042 G; 1012 T; 0 U; 0 Other; | |
| SQ | | |
| Query Match 92.2%; Score 903; DB 8; Length 3975; | | |
| Best Local Similarity 96.6%; Pred. No. 2.3e-258; | | |
| Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1; | | |
| Qy | 1 ACCAAGCATTTACTGTTATTATCAACCCGTTTGGAGGAAAGGACAGGCGGAT 60 | |
| Db | 24 ACCAAGCATTTACTGTTATTATCAACCCGTTTGGAGGAAAGGACAGGCGGAT 83 | |
| Qy | 61 ATATGAAGAAAGTGGCACCACCTGTTCACTTAGCCTCCATCACCCTGACATCATCGG 120 | |
| Db | 84 ATATGAAGAAAGTGGCACCACCTGTTCACTTAGCCTCCATCACCCTGACATCATCG- 142 | |
| Qy | 121 TAACAAATTCATGTTAACTATGTAGAAAGTAATTACTGAACATGCTAATCAGGCCAAGGA 180 | |
| Db | 143 -----TTACTGAACATGCTAATCAGGCCAAGGA 170 | |
| Qy | 181 GACTCTGTATGATTAACATAGACAAATACAGCGGCATCTGTGTGCGCGGAGATGG 240 | |
| Db | 171 GACTCTGTATGATTAACATAGACAAATACAGCGGCATCTGTGTGCGCGGAGATGG 230 | |
| Qy | 241 TATGTTACGAGGTCGTCACGGTCTCATTTGGAGGACGAGAGCGCGGGTCTGA 300 | |
| Db | 231 TATGTTACGAGGTCGTCACGGTCTCATTTGGAGGACGAGAGCGCGGGTCTGA 290 | |
| Qy | 301 CCAGAACACCCCGGGTGTGCTGCTCCAGTAGCCTCCCGATTGGAATCATTTCCCGC 360 | |
| Db | 291 CCAGAACACCCCGGGTGTGCTGCTCCAGTAGCCTCCCGATTGGAATCATTTCCCGC 350 | |
| Qy | 361 AGGGTCAACGACTGCGTGTGTTACTCCAGTGGGACGAGCGACGAGAACTCGGC 420 | |
| Db | 351 AGGGTCAACGACTGCGTGTGTTACTCCAGTGGGACGAGCGACGAGAACTCGGC 410 | |
| Qy | 421 GCTCATATCTGTTGGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACACAG 480 | |
| Db | 411 GCTCATATCTGTTGGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACACAG 470 | |
| Qy | 481 CACACTCTTCCTACTCCGTCCTGCTGGCTTACCGGCTTCTACGGGACATCATCAA 540 | |
| Db | 471 CACACTCTTCCTACTCCGTCCTGCTGGCTTACCGGCTTCTACGGGACATCATCAA 530 | |
| Qy | 541 GCAGGTAGAGAAACGGTGGTGGGTCCTTCCAGATACGACTTTTCAGGTTTAAAGAC 600 | |
| Db | 531 GCAGGTAGAGAAACGGTGGTGGGTCCTTCCAGATACGACTTTTCAGGTTTAAAGAC 590 | |
| Qy | 601 CTTCCTCTCCACCACTCTATGAGGACAGTGTCTTCTCTCCCTGCACAAACACAGGT 660 | |
| Db | 591 CTTCCTCTCCACCACTCTATGAGGACAGTGTCTTCTCTCCCTGCACAAACACAGGT 650 | |
| Qy | 661 GGGATCTCCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCGAAAGCAA 720 | |
| Db | 651 GGGATCTCCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCGAAAGCAA 710 | |
| Qy | 721 GCAGCAGCTGGAGGAGGACGAAGAAAGCACTGTATGTTGGAAAGCTGCGGAGGACGT 780 | |
| Db | 711 GCAGCAGCTGGAGGAGGACGAAGAAAGCACTGTATGTTGGAAAGCTGCGGAGGACGT 770 | |
| Qy | 781 GGAGGAGTGGCAAGTCTGTGTGGAAAGTTCCTGGCCATCAATGCCACAAACATGTCTCTG 840 | |
| Db | 771 GGAGGAGTGGCAAGTCTGTGTGGAAAGTTCCTGGCCATCAATGCCACAAACATGTCTCTG 830 | |
| Qy | 841 TGCTTGTGCGCGGAGCCCCAGGGGCTCTCCCCGGCTGCCACATTGGGAGACGGGTCTTC 900 | |
| Db | 831 TGCTTGTGCGCGGAGCCCCAGGGGCTCTCCCCGGCTGCCACATTGGGAGACGGGTCTTC 890 | |
| Qy | 901 TGACCTCATCTCTATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 960 | |
| Db | 891 TGACCTCATCTCTATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 950 | |
| Qy | 961 CACCAACAGCAGGACCCAG 979 | |
| Db | 951 CACCAACAGCAGGACCCAG 969 | |
| RESULT 5 | | |
| AAA50510 | | |
| ID | AAA50510 standard; cDNA; 4231 BP. | |
| XX | AC AAA50510; | |
| DT | 05-DEC-2000 (first entry) | |
| XX | Human sphingosine kinase C cDNA. | |
| XX | Sphingosine kinase C; SKC; human; drug screening; infection; | |
| KW | antiinflammatory; antiallergic; anticancer; inflammation; allergy; | |
| KW | cancer; therapy; diagnosis; ds. | |
| XX | Homo sapiens. | |
| XX | Key Location/Qualifiers | |
| FH | CDS 71..1453 | |
| FT | /*tag= a | |
| FT | WO200052173-A2. | |
| XX | 08-SEP-2000. | |
| XX | 02-MAR-2000; 2000WO-CA000223. | |
| XX | 02-MAR-1999; 99US-0122516P. | |
| XX | (ALLX) NPS ALLELIX CORP. | |
| XX | Munroe D, Gupta A, Falzone GR; | |
| XX | WPI; 2000-572185/53. | |
| XX | P-PSDB; AAY96059. | |
| XX | New human sphingosine kinase A, B and C polynucleotides and polypeptides | |
| XX | useful in e.g. chromosome and gene mapping, and detecting inflammation or | |
| XX | disease associated with abnormal levels of sphingosine kinase expression. | |
| XX | Disclosure; Fig 7; 81pp; English. | |
| XX | The present sequence is that of an isolated polynucleotide encoding human | |
| CC | sphingosine kinase C (SKC, see AAY96059), an enzyme that phosphorylates | |
| CC | sphingosine to form sphingosine 1-phosphate. The polynucleotide was | |
| CC | isolated from an HeLa cDNA library by PCR amplification. The invention | |
| CC | provides polynucleotides (see AAA50508-10) and polypeptides (see AAY96057 | |
| CC | -59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC. | |
| CC | The polynucleotides may be used as hybridization probes, in the | |
| CC | construction of PCR primers for chromosome and gene mapping, in the | |
| CC | recombinant production of SKA, SKB and SKC, and in the generation of | |
| CC | antisense DNA or RNA. They can be used to detect inflammation or disease | |
| CC | associated with abnormal levels of SK expression, or to detect | |
| CC | differences in gene sequence between normal and carrier or affected | |

CC individuals. Host cells expressing SK can be used in drug screening.
CC Human SK specific antibodies, inhibitors, ligands or their analogues are
CC useful as bioactive agents to treat inflammation or disease including
CC viral, bacterial or fungal infections, allergic responses, mechanical
CC injury associated with trauma, hereditary diseases, lymphoma or
CC carcinoma, and other conditions with activate the genes of kidney, lung,
CC heart, lymphoid or tissues of the nervous system
XX
SQ Sequence 4231 BP; 1022 A; 1021 C; 1086 G; 1067 T; 0 U; 35 Other;

Query Match 92.2%; Score 903; DB 3; Length 4231;
Best Local Similarity 96.6%; Pred. No. 2.4e-258;
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;
QY 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGGACAGCGCGAT 60
DB |||||||
QY 226 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGGACAGCGCGAT 285
DB |||||||
QY 61 ATATGAAAGAAAGTGGCACCACCTGTTTCACTTAGCCCTTCCATCACCCTGACATCATCGG 120
DB |||||||
QY 286 ATATGAAAGAAAGTGGCACCACCTGTTTCACTTAGCCCTTCCATCACCCTGACATCATCG- 344
DB |||||||
QY 121 TAACAAATTCATGTTAACTAGTATAGAAATTAATCTGAAACATGTAATCAAGGCAAGGA 180
DB |||||||
QY 345 -----TTACTGAAACATGCTAATCAGGCAAGGA 372
DB |||||||
QY 181 GACTCTGATGAGATTACATAGACAATACGCGGCATCTCTGTGCGCGGAGATGG 240
DB |||||||
QY 373 GACTCTGATGAGATTACATAGACAATACGCGGCATCTCTGTGCGCGGAGATGG 432
DB |||||||
QY 241 TATGTTACGAGGTGCTGCAAGTCTGATTTGGAGGACGAGAGCGCGCGGTCGA 300
DB |||||||
QY 433 TATGTTACGAGGTGCTGCAAGTCTGATTTGGAGGACGAGAGCGCGCGGTCGA 492
DB |||||||
QY 301 CCAGAACCAACCCCGGGTGTGTGTCCTCCAGTAGCCTCCGATGCGAATGCAATTCCTCGC 360
DB |||||||
QY 493 CCAGAACCAACCCCGGGTGTGTGTCCTCCAGTAGCCTCCGATGCGAATGCAATTCCTCGC 552
DB |||||||
QY 361 AGGGTCAACGAGTCTGCTGTTACTCCACCGTGGGACGAGCGGCAAGAACTCGG 420
DB |||||||
QY 553 AGGGTCAACGAGTCTGCTGTTACTCCACCGTGGGACGAGCGGCAAGAACTCGG 612
DB |||||||
QY 421 GCTGCATATCTGTTGGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACACAG 480
DB |||||||
QY 613 GCTGCATATCTGTTGGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACACAG 672
DB |||||||
QY 481 CACACTCTTCTGCTACTCCGTTGCTGCTGGGCTACGGCTTCTACGGGACATCAAA 540
DB |||||||
QY 673 CACACTCTTCTGCTACTCCGTTGCTGCTGGGCTACGGCTTCTACGGGACATCAAA 732
DB |||||||
QY 541 GGACAGTCAGAGAAACGGTGGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 600
DB |||||||
QY 733 GGACAGTCAGAGAAACGGTGGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 792
DB |||||||
QY 601 CTTCTCTCCCAACCACTCTATGAAGGACAGTGTCTTCTCCTCCCTGCACACACAGGT 660
DB |||||||
QY 793 CTTCTCTCCCAACCACTCTATGAAGGACAGTGTCTTCTCCTCCCTGCACACACAGGT 852
DB |||||||
QY 661 GGGATCTCCAAAGGATAGAAAGCCCTGCGGCGAGGATGCTTTGTCAGGCAAGCA 720
DB |||||||
QY 853 GGGATCTCCAAAGGATAGAAAGCCCTGCGGCGAGGATGCTTTGTCAGGCAAGCA 912
DB |||||||
QY 721 GCAGCAGCTGGAGGAGGACAGAAAGACTGTATGTTTGAAGCTGCGGAGGAGCT 780
DB |||||||
QY 913 GCAGCAGCTGGAGGAGGACAGAAAGACTGTATGTTTGAAGCTGCGGAGGAGCT 972
DB |||||||
QY 781 GGAGGAGTGGCAAGTCTGTGTTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTG 840
DB |||||||
QY 973 GGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTG 1032
DB |||||||
QY 841 TGCTTGTGCGCGGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGTCTTC 900
DB |||||||
QY 1033 TGCTTGTGCGCGGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGTCTTC 1092
DB |||||||

QY 901 TGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGCA 960
DB |||||||
QY 1093 TGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGCA 1152
DB |||||||
QY 961 CACCAACCAGCAGGACCAG 979
DB |||||||
QY 1153 CACCAACCAGCAGGACCAG 1171
DB |||||||
RESULT 6
ABL40828
ID ABL40828 standard; cDNA; 4413 BP.
XX ABL40828;
AC ABL40828;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human sphingosine kinase-like protein encoding cDNA.
XX
KW Human sphingosine kinase-like protein; intracellular signalling; gene;
cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1689
FT /*tag= a
FT /product= "sphingosine kinase-like protein"
FT /note= "see ABB07857"
FT CDS 76..1689
FT /*tag= b
FT /product= "sphingosine kinase-like protein"
FT /note= "see ABB07856"
XX
PN W0200228906-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-EP011516.
XX
PR 06-OCT-2000; 2000US-0238005P.
PR 23-AUG-2001; 2001US-0314113P.
XX
PA (FARB) BAYER AG.
XX
PI Kossida S, Encinas J;
XX WPI; 2002-340094/37.
XX P-PSDB; ABB07856, ABB07857.
DR
DR
XX
XX
PT New reagent for modulating the activity of sphingosine kinase-like
protein polypeptide or polynucleotide and treating cancer, asthma,
PT allergy, an autoimmune disease, or a central or peripheral nervous system
PT disorder.
PT
PT
XX
PS Claim 1; Fig 9; 120pp; English.
XX
XX
CC The invention relates to a human sphingosine kinase-like protein. The
polypeptide can be expressed by standard recombinant methodology. The
CC sphingosine kinase-like protein and gene can be used to regulate
intracellular signalling and consequently cell proliferation and
CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
CC peripheral nervous system disorders (e.g. Parkinson's disease). The
CC present sequence represents the human sphingosine kinase-like protein
encoding cDNA
XX
SQ Sequence 4413 BP; 1020 A; 1107 C; 1207 G; 1079 T; 0 U; 0 Other;

Query Match 92.2%; Score 903; DB 6; Length 4413;
Best Local Similarity 96.6%; Pred. No. 2.5e-258;

| | |
|---------------------------------------------------------------|----------------------------------------------------------------------|
| Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1; | |
| QY 1 | ACCAAAGCATTTACTGGTATTATCAACCCGTTTCGAGGAAAGACAAGCAAGCGGAT 60 |
| Db | |
| 462 | ACCAAAGCATTTACTGGTATTATCAACCCGTTTCGAGGAAAGACAAGCAAGCGGAT 521 |
| QY | 61 ATATCAAAAGAAAGTGGACCACTGTTACCTTAGCCTCCATCACCACCTGACATCATCGG 120 |
| Db | |
| 522 | ATATCAAAAGAAAGTGGACCACTGTTACCTTAGCCTCCATCACCACCTGACATCATCG- 580 |
| QY | 121 TAACAAATTCATGTTAACTATGTAGAAGTAATTAAGCAATGCTATCAAGGCCAAGGA 180 |
| Db | |
| 581 | -----TTACTGAACATGCTATCAGGCCACAGGA 608 |
| QY | 181 GACTCTGATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGG 240 |
| Db | |
| 609 | GACTCTGATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGG 668 |
| QY | 241 TATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGGCGCGGGGTGCA 300 |
| Db | |
| 669 | TATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGGTGCA 728 |
| QY | 301 CCAGAACCAACCCCGGCTGCTGCTGCCAGTAGCTCCGAGATTGGAATCATTTCCCGC 360 |
| Db | |
| 729 | CCAGAACCAACCCCGGCTGCTGCTGCCAGTAGCTCCGAGATTGGAATCATTTCCCGC 788 |
| QY | 361 AGGTCACACGCACTCGTGTGTTACTCCACCGTGGGACCAAGCGACGAGAAACCTCGGC 420 |
| Db | |
| 789 | AGGTCACACGCACTCGTGTGTTACTCCACCGTGGGACCAAGCGACGAGAAACCTCGGC 848 |
| QY | 421 GCTGCATATCTGTTGGGACTCCTGGCCATGATGTCTCAGTCCACCAACAG 480 |
| Db | |
| 849 | GCTGCATATCTGTTGGGACTCCTGGCCATGATGTCTCAGTCCACCAACAG 908 |
| QY | 481 CACACTCTCTGCTACTCTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAA 540 |
| Db | |
| 909 | CACACTCTCTGCTACTCTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAA 968 |
| QY | 541 GGACAGTGAAGAAACGGTGTGGGTCTTTCGACAGATACGACTTTTTCAGGTTTAAAGAC 600 |
| Db | |
| 969 | GGACAGTGAAGAAACGGTGTGGGTCTTTCGACAGATACGACTTTTTCAGGTTTAAAGAC 1028 |
| QY | 601 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTGTCACACACACGGT 660 |
| Db | |
| 1029 | CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTGTCACACACACGGT 1088 |
| QY | 661 GGGATCTCCAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGSCAAAGCAA 720 |
| Db | |
| 1089 | GGGATCTCCAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGSCAAAGCAA 1148 |
| QY | 721 GCAGCAGCTGGAGGAGGAGCAAGAAAGCACTGTATGTTTGGAAAGCTGGGGAGGACGT 780 |
| Db | |
| 1149 | GCAGCAGCTGGAGGAGGAGGAGCAAGAAAGCACTGTATGTTTGGAAAGCTGGGGAGGACGT 1208 |
| QY | 781 GGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTG 840 |
| Db | |
| 1209 | GGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTG 1268 |
| QY | 841 TGCTTGTGCGCGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 900 |
| Db | |
| 1269 | TGCTTGTGCGCGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 1328 |
| QY | 901 TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTCGAGATTCTCATCAGGCA 960 |
| Db | |
| 1329 | TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTCGAGATTCTCATCAGGCA 1388 |
| QY | 961 CACCAACAGGAGGACCAG 979 |
| Db | |
| 1389 | CACCAACAGGAGGACCAG 1407 |

RESULT 7
ADJ96598

| | |
|-----|---------------------------------------------------------------------------|
| ID | ADJ96598 standard; DNA; 4429 BP. |
| XX | |
| AC | ADJ96598; |
| XX | |
| DT | 06-MAY-2004 (first entry) |
| XX | |
| DE | Human lipid kinase KIAA1646 DNA SeqID 55. |
| XX | |
| KW | gene; ds; kinase; human; SNP; single nucleotide polymorphism; |
| KW | tyrosine protein kinase; serine/threonine protein kinase; PKK; STK; |
| KW | gene therapy; cancer; immune-related disease; cardiovascular disease; |
| KW | brain; neuronal associated disease; metabolic; inflammatory disorder; |
| KW | cystostatic; neuroprotective; immunomodulator; antinflammatory; |
| XX | lipid kinase; KIAA1646. |
| XX | |
| OS | Homo sapiens. |
| 39. | |
| XX | |
| FH | Key Location/Qualifiers |
| FT | variation replace(2391,g) |
| FT | /tag= a |
| FT | /standard_name= "Single nucleotide polymorphism" |
| FT | variation replace(2577,g) |
| FT | /tag= b |
| FT | /standard_name= "Single nucleotide polymorphism" |
| FT | variation replace(3020,c) |
| FT | /tag= c |
| FT | /standard_name= "Single nucleotide polymorphism" |
| FT | variation replace(3769,c) |
| FT | /tag= d |
| FT | /standard_name= "Single nucleotide polymorphism" |
| FT | variation replace(4272,g) |
| FT | /tag= e |
| FT | /standard_name= "Single nucleotide polymorphism" |
| XX | |
| PN | WO2004006838-A2. |
| XX | |
| PD | 22-JAN-2004. |
| XX | |
| PF | 15-JUL-2003; 2003WO-US021730. |
| XX | |
| PR | 15-JUL-2002; 2002US-0395632P. |
| XX | |
| PA | (SUGE-) SUGEN INC. |
| XX | |
| PI | Whyte D, Manning G, Caenepeel S; |
| XX | |
| DR | WPI; 2004-122753/12. |
| XX | |
| DR | P-PSDB; ADJ96664. |
| XX | |
| XX | New nucleic acid molecule encoding a kinase polypeptide, useful for |
| PT | preparing a composition for treating diseases or disorders, e.g., cancer, |
| PT | or neurological, immunological or inflammatory disorders. |
| XX | |
| PS | Example 1; SEQ ID NO 55; 366pp; English. |
| XX | |
| CC | This invention relates to a novel isolated, enriched or purified nucleic |
| CC | acid molecule that encodes a kinase polypeptide. Specifically, it relates |
| CC | to human tyrosine and serine/threonine protein kinases (PTK's and STK's), |
| CC | as well as protein kinase-like enzymes. The present invention describes |
| CC | screening methods to identify agonists, antagonists and antibodies that |
| CC | can be used to modulate the activity or function of the mammalian kinase |
| CC | enzymes. As such, these compositions can be used for gene therapy |
| CC | purposes to treat diseases or disorders including cancer, immune-related |
| CC | diseases, cardiovascular disease, brain or neuronal associated disease, |
| CC | metabolic and inflammatory disorders. Accordingly, they exhibit |
| CC | cytostatic, neuroprotective, immunomodulator and antinflammatory |
| CC | activities. This polynucleotide sequence is a human kinase DNA sequence |
| CC | of the invention. |
| XX | |
| SQ | Sequence 4429 BP; 1020 A; 1115 C; 1213 G; 1081 T; 0 U; 0 Other; |

Query Match

92.2%; Score 903; DB 12; Length 4429;

Best Local Similarity 96.6%; Pred. No. 2.5e-258;
Matches 946; Conservative 0; Mismatches 0;

| | | | |
|----|------|------------------------------------------------------------------|------|
| Qy | 1 | ACAAAGCATTTTACTGGTATTTTATCAACCGCTTTGGAGGAAAAGCACAAGCAAGCCGAT | 60 |
| Db | 478 | ACCAAGCATTTTACTGGTATTTTATCAACCGCTTTGGAGGAAAAGCACAAGCAAGCCGAT | 537 |
| Qy | 61 | ATATGAAGAAAAGTGGCACCACATGTTCAACCTTAGCCTCCATCAACCACTGCATCATCGG | 120 |
| Db | 538 | ATATGAAGAAAAGTGGCACCACATGTTCAACCTTAGCCTCCATCAACCACTGCATCATCG | 596 |
| Qy | 121 | TAACAAAATTTCTATGTTAACTATGTAGAAGTAATTTACTCAACATGCTTAATCAGGCCAAGGA | 180 |
| Db | 597 | -----TTTACTGAACATGCTTAATCAGGCCAAGGA | 624 |
| Qy | 181 | GACTCTGTATGAGATTAACTAGACAATAACGACGGCATCGTCTGTCTGGCGGAGATGG | 240 |
| Db | 625 | GACTCTGTATGAGATTAACTAGACAATAACGACGGCATCGTCTGTCTGGCGGAGATGG | 684 |
| Qy | 241 | TATGTTCAAGCAGAGTGTGACCGTCTGATTGGGAGGACGACAGGAGAGCGCCGGGTCTGA | 300 |
| Db | 685 | TATGTTCAAGCAGAGTGTGACCGTCTGATTGGGAGGACGACAGGAGAGCGCCGGGTCTGA | 744 |
| Qy | 301 | CCAGAACACCCCGGGGTGTCGTGGTCCCCAGTAGCTCTCGGATTTGGAATTCATTCGGC | 360 |
| Db | 745 | CCAGAACACCCCGGGGTGTCGTGGTCCCCAGTAGCTCTCGGATTTGGAATTCATTCGGC | 804 |
| Qy | 361 | AGGGTCAACGACGTCGTGTTTACTCCACCGTGGGACACGACGACGAGAAACCTCGGC | 420 |
| Db | 805 | AGGGTCAACGACGTCGTGTTTACTCCACCGTGGGACACGACGACGAGAAACCTCGGC | 864 |
| Qy | 421 | GCTGCATATCGTTGTGGGACTCGCTGGCCATGGATGTGTCCTCAGTCCACCAACAACAG | 480 |
| Db | 865 | GCTGCATATCGTTGTGGGACTCGCTGGCCATGGATGTGTCCTCAGTCCACCAACAACAG | 924 |
| Qy | 481 | CACACTCTTCGCTACTCCGTCCTCTGCTGGGTCTACGGCTTTCACGGGGACATCATCAA | 540 |
| Db | 925 | CACACTCTTCGCTACTCCGTCCTCTGCTGGGTCTACGGCTTTCACGGGGACATCATCAA | 984 |
| Qy | 541 | GGACAGTGAGAAGAAACGGTGGTGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGAC | 600 |
| Db | 985 | GGACAGTGAGAAGAAACGGTGGTGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGAC | 1044 |
| Qy | 601 | CTTCCTCTCCCAACACTGCTATGAAGGACAGTGTCTTCCTCCCTGCACAACAACCGGT | 660 |
| Db | 1045 | CTTCCTCTCCCAACACTGCTATGAAGGACAGTGTCTTCCTCCCTGCACAACAACCGGT | 1104 |
| Qy | 661 | GGGATCTCAAGGATAGGAAGCCCTGCCGGGCAAGATGCTTTGTTTGCAGGCAAGCAA | 720 |
| Db | 1105 | GGGATCTCAAGGATAGGAAGCCCTGCCGGGCAAGATGCTTTGTTTGCAGGCAAGCAA | 1164 |
| Qy | 721 | GCACAGCTGAGGAGGACGAGAAGAAGACTGTATGGTTTGAAGCTGCGGAGGACGT | 780 |
| Db | 1165 | GCACAGCTGAGGAGGAGCAGAAGAAGACTGTATGGTTTGAAGCTGCGGAGGACGT | 1224 |
| Qy | 781 | GGAGGAGTGGCAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTG | 840 |
| Db | 1225 | GGAGGAGTGGCAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTG | 1284 |
| Qy | 841 | TGCTTTGCGCGGAGCCCCAGGGGCCTCTCCCGGCTGCCACTTGGGAGCGGTCCTC | 900 |
| Db | 1285 | TGCTTTGCGCGGAGCCCCAGGGGCCTCTCCCGGCTGCCACTTGGGAGCGGTCCTC | 1344 |
| Qy | 901 | TGACCTCATCTCATCCGGAATGCTCCAGGTTCAAATTTTCTGAGATTTCTCATCAGGCA | 960 |
| Db | 1345 | TGACCTCATCTCATCCGGAATGCTCCAGGTTCAAATTTTCTGAGATTTCTCATCAGGCA | 1404 |
| Qy | 961 | CACCAACGACGAGACCAG | 979 |
| Db | 1405 | CACCAACGACGAGACCAG | 1423 |

ADP55247
ID ADP55247 standard; cDNA; 4445 BP.
XX
XX AC ADP55247;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human PRO cDNA sequence SEQ ID NO:1223
XX
XX KW human; PRO; immune related disease; inf
KW immune response stimulation; antiallerg
KW antiasthmatic; antidiabetic; antiinfla
KW antirheumatic; antithyroid; CNS; derma
KW haemostatic; hepatotropic; immunostimu
KW nephrotropic; neuroprotective; osteopa
KW virucide; gene therapy; gene; ss.
XX
XX Homo sapiens.
OS
XX WO200403956-A2.
PN
XX 13-MAY-2004.
PD
XX 28-OCT-2003; 2003WO-US034381.
PF
XX 29-OCT-2002; 2002US-0422472P.
PR
XX (GETH) GENENTECH INC.
PA
XX Aggarwal S, Clark H, Gurney AL, Scho
PI Wood WI, Wu TD;
XX
XX WPI: 2004-376182/35.
DR P-FSDB; ADP55248.
DR
XX New PRO polynucleotides and polypeptid
PT and treating an immune related diseas
PT erythematosus, rheumatoid arthritis, d
PT stimulating an immune response.
XX
XX Claim 2; SEQ ID NO 1223; 3009pp; Englis
PS
XX The present invention describes an isol
CC described: (1) a vector comprising (1)
CC vector of (1); (3) a process for produ
CC isolated PRO polypeptide; (5) a chimer
CC polypeptide of (4) fused to a heterolog
CC antibody which specifically binds to a
CC composition of matter comprising a poly
CC antagonist of the polypeptide or an ant
CC polypeptide in combination with a carri
CC comprising a container, a label on the
CC matter of (7); (9) a method of treatin
CC mammal; (10) a method for determining t
CC in a sample suspected of having the po
CC diagnosing an immune related disease o
CC in mammal; (12) a method of identifyi
CC mimics the activity of or expressing th
CC ; and (13) a method of stimulating the
CC PRO sequences have anti-allergic, antia
CC antiasthmatic, antidiabetic, antiinfla
CC antirheumatic, antithyroid, CNS, derma
CC haemostatic, hepatotropic, immunostimu
CC nephrotropic, neuroprotective, osteopa
CC virucide activities, and can be used i
CC (I) and the encoded polypeptides, compo
CC useful in diagnosing and treating an im
CC stimulating an immune response. The pro
CC PRO nucleotide sequence from the prese
XX Sequence 4445 BP; 1018 A; 1134 C; 1216

RESULT 8

Query Match

92.2%; Score 903; DB 13; Length 4445;

| | | | |
|--------------------------------------------------|------|---------------------------------------------------------------|------|
| Best Local Similarity 96.6%; Pred. No. 2.5e-258; | | Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1; | |
| QY | 1 | ACCAAGCACTTTACTGGTATTTATCAACCCGTTTGGAGAAAGGACAAAGGCAAGCGGAT | 60 |
| DB | 510 | ACCAAGCACTTTACTGGTATTTATCAACCCGTTTGGAGAAAGGACAAAGGCAAGCGGAT | 569 |
| QY | 61 | ATATGAAGAAAGTGGACCACTGTTTACCTTAGCTCCATCACCCTGATCATCGG | 120 |
| DB | 570 | ATATGAAGAAAGTGGACCACTGTTTACCTTAGCTCCATCACCCTGATCATCGG | 628 |
| QY | 121 | TAACAAATTCATGTTAACTATGTAGAGTAATTAAGCAATGCTAATCAGGCCAAGGA | 180 |
| DB | 629 | -----TTACTGAACATGCTAATCAGGCCAAGGA | 656 |
| QY | 181 | GACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGCGGAGATGG | 240 |
| DB | 657 | GACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGCGGAGATGG | 716 |
| QY | 241 | TATGTTTCAGCGAGGTGTCGACCGTCTGANTGGGAGGACGACAGAGGCGCGGGGTGA | 300 |
| DB | 717 | TATGTTTCAGCGAGGTGTCGACCGTCTGANTGGGAGGACGACAGAGGCGCGGGGTGA | 776 |
| QY | 301 | CCAGAACACACCCCGGGTGTCTGTTACTCCACCGTGGGACACGACGAGCAAACTCCGC | 360 |
| DB | 777 | CCAGAACACACCCCGGGTGTCTGTTACTCCACCGTGGGACACGACGAGCAAACTCCGC | 836 |
| QY | 361 | AGGGTCAACGGAAGTGGTGTCTGTTACTCCACCGTGGGACACGACGAGCAAACTCCGC | 420 |
| DB | 837 | AGGGTCAACGGAAGTGGTGTCTGTTACTCCACCGTGGGACACGACGAGCAAACTCCGC | 896 |
| QY | 421 | GCTCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAG | 480 |
| DB | 897 | GCTCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAG | 956 |
| QY | 481 | CACACTCTCTGCTACTCGTGTCTGCTGGGCTACGGCTTCTACGGGACATCAAA | 540 |
| DB | 957 | CACACTCTCTGCTACTCGTGTCTGCTGGGCTACGGCTTCTACGGGACATCAAA | 1016 |
| QY | 541 | GGACAGTGAGAAAGAAACGGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGAC | 600 |
| DB | 1017 | GGACAGTGAGAAAGAAACGGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGAC | 1076 |
| QY | 601 | CTTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCTGCTGACACACACCGT | 660 |
| DB | 1077 | CTTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCTGCTGACACACACCGT | 1136 |
| QY | 661 | GGGATCTCCAGGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCAAGCAA | 720 |
| DB | 1137 | GGGATCTCCAGGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCAAGCAA | 1196 |
| QY | 721 | GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGGAGT | 780 |
| DB | 1197 | GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGGAGT | 1256 |
| QY | 781 | GGAGAGTGGCAAGTCGTCTGTGGAAAGTTTCTGGCCATCAATGCCAACAAATGCTCTG | 840 |
| DB | 1257 | GGAGAGTGGCAAGTCGTCTGTGGAAAGTTTCTGGCCATCAATGCCAACAAATGCTCTG | 1316 |
| QY | 841 | TGCTTGTGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTC | 900 |
| DB | 1317 | TGCTTGTGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTC | 1376 |
| QY | 901 | TGACCTCATCTTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGCA | 960 |
| DB | 1377 | TGACCTCATCTTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGCA | 1436 |
| QY | 961 | CACCAACCAAGGAGGACAG 979 | |
| DB | 1437 | CACCAACCAAGGAGGACAG 1455 | |

| | | |
|----------|-----------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------|
| ABX70921 | ID | ABX70921 standard; cDNA; 4432 BP. |
| XX | AC | ABX70921; |
| XX | DT | 05-MAR-2003 (first entry) |
| XX | XX | Novel human cDNA sequence #146. |
| DE | Human; gene; ss; nervous system disorder; peripheral neuropathy; | |
| XX | Human; gene; ss; nervous system disorder; peripheral neuropathy; | |
| XX | Huntington's disease; amyotrophic lateral sclerosis; haemophilia; | |
| KW | neurodegenerative disease; Parkinson's disease; Alzheimer's disease; | |
| KW | autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis; | |
| KW | insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound; | |
| KW | ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke; | |
| KW | fibrosis; reperfusion injury; infection; allergic rhinitis; asthma; | |
| KW | coagulation disorder; cancer; tumour; inflammatory disease; septic shock; | |
| KW | Crohn's disease; anaphylaxis; proliferation; chemotactic; chemokinetic; | |
| KW | differentiation; stem cell growth factor; haematopoiesis; haemostatic; antiinflammatory; expressed sequence tag; EST. | |
| XX | OS | Homo sapiens. |
| XX | XX | WO20020281731-A2. |
| PN | XX | 17-OCT-2002. |
| PD | XX | 29-JAN-2002; 2002WO-US001222. |
| PF | XX | 30-JAN-2001; 2001US-00774528. |
| PR | XX | (HYSE-) HYSEQ INC. |
| XX | PA | (GOOD/) GOODRICH R W. |
| XX | PA | Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F; |
| PI | XX | Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Dmanac RT; |
| XX | XX | WPI; 2003-058563/05. |
| DR | XX | Novel polypeptide useful for treating neurodegenerative diseases, myeloid |
| XX | PT | or lymphoid cell disorders, bone disorders, mechanical and traumatic |
| PT | PT | disorders, coagulation disorders, and inflammatory diseases. |
| XX | XX | Claim 1; Page; 612pp; English. |
| PS | XX | This invention relates to the cDNA sequences encoding an isolated novel |
| CC | CC | human polypeptide. The protein encoded by the nucleic acid of the |
| CC | CC | invention is useful for treating central and peripheral nervous system |
| CC | CC | diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic |
| CC | CC | lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease, |
| CC | CC | Alzheimer's disease); autoimmune disease (e.g. systemic lupus |
| CC | CC | erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus) |
| CC | CC | ; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopaenia) |
| CC | CC | ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis, |
| CC | CC | osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head |
| CC | CC | trauma); lung or liver fibrosis; reperfusion injury in various tissues; |
| CC | CC | bacterial, viral or fungal infections; allergic conditions such as |
| CC | CC | allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia); |
| CC | CC | cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's |
| CC | CC | disease, anaphylaxis). The protein may be used to inhibit the growth, |
| CC | CC | infection or function of infectious agents such as bacteria, fungi, |
| CC | CC | viruses, or to effect bodily characteristics, biorhythms or circadian |
| CC | CC | cycles of rhythms. The protein may also have |
| CC | CC | proliferation/differentiation, stem cell growth factor, haematopoiesis |
| CC | CC | regulation, immune stimulating or suppressing, chemotactic/chemokinetic, |
| CC | CC | haemostatic and thrombolytic, receptor/ligand, and antiinflammatory |
| CC | CC | activities. The cDNA sequences of the invention are useful for expressing |
| CC | CC | recombinant protein for analysis. The present sequence represents a novel |
| CC | CC | human cDNA sequence of the invention, this sequence is an expressed |
| CC | CC | sequence tag (EST) and was identified using subtractive hybridisation |
| XX | XX | Sequence 4432 BP; 970 A; 1122 C; 1282 G; 1058 T; 0 U; 0 Other; |

```
Query Match      92.1%; Score 901.4; DB 8; Length 4432;
Best Local Similarity 96.5%; Pred. No. 7.5e-258;
Matches 945; Conservative 0; Mismatches 1; Indels 33; Gaps 1;

QY 1 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAAGACAAAGCAAGCGGAT 60
DB 1590 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAAGACAAAGCAAGCGGAT 1649

QY 61 ATATGAAAGAAAGTGGCACCACCTGTTCCACCTTAGCTTACCTCCATCACCCTGACATCATCGG 120
DB 1650 ATATGAAAGAAAGTGGCACCACCTGTTCCACCTTAGCTTACCTCCATCACCCTGACATCATCG- 1708

QY 121 TAACAAATTCATGTTAACTAGTAGAAGTAATTAAGCAATGCTAATCAGGCAAGGA 180
DB 1709 -----TTACTGAACATGCTAATCAGGCAAGGA 1736

QY 181 GACTCTGTATGAGATTAAACATAGACAAATATACAGCGGCATCGTCTGTGTCGGCGAGATGG 240
DB 1737 GACTCTGTATGAGATTAAACATAGACAAATATACAGCGGCATCGTCTGTGTCGGCGAGATGG 1796

QY 241 TATGTTTACGCGAGTGTCTGCGGTTCTGATTTGGGAGGACGAGAGGCGCGGGGTCTGA 300
DB 1797 TATGTTTACGCGAGTGTCTGCGGTTCTGATTTGGGAGGACGAGAGGCGCGGGGTCTGA 1856

QY 301 CCAGAACACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGGATTGGAATCATTTCCCGC 360
DB 1857 CCAGAACACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGGATTGGAATCATTTCCCGC 1916

QY 361 AGGTCAACGAGCTGCTGTGTTACTCCACGTTGGGACACGAGCGGACGAAACCTTGGC 420
DB 1917 AGGTCAACGAGCTGCTGTGTTACTCCACGTTGGGACACGAGCGGACGAAACCTTGGC 1976

QY 421 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAG 480
DB 1977 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAG 2036

QY 481 CACACTCTTCGCTACTCCGTTCCCTGCTGGGCTTACGGGCTTCTACGGGACATCATCA 540
DB 2037 CACACTCTTCGCTACTCCGTTCCCTGCTGGGCTTACGGGCTTCTACGGGACATCATCA 2096

QY 541 GGACAGTGAAGAAACGGTGGTGGTCTTGGCCAGATACACTTTTCAGTTTAAAGAC 600
DB 2097 GGACAGTGAAGAAACGGTGGTGGTCTTGGCCAGATACACTTTTCAGTTTAAAGAC 2156

QY 601 CTTCTCTCCACCACTCTATGAAGGACAGTGTCTTCTCTCTGACCAACACACGGT 660
DB 2157 CTTCTCTCCACCACTCTATGAAGGACAGTGTCTTCTCTCTGACCAACACACGGT 2216

QY 661 GGGATCTCCAAAGGATAGGAAGCCCTCGCGGCGAGGATGCTTTGTTGCGAGGCAAGCAA 720
DB 2217 GGGATCTCCAAAGGATAGGAAGCCCTCGCGGCGAGGATGCTTTGTTGCGAGGCAAGCAA 2276

QY 721 GCAGAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAGTGGAGTGGGAGGACGT 780
DB 2277 GCAGAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAGTGGGAGGACGT 2336

QY 781 GGAGAGTGGCAAGTCTGTGGAAGTCTTCTGSCCATCAATGCCACAAACATGTCTCTG 840
DB 2337 GGAGAGTGGCAAGTCTGTGGAAGTCTTCTGSCCATCAATGCCACAAACATGTCTCTG 2396

QY 841 TGCTTGTCCCGGAGCCCCAGGGGCTCTCCCGGCTGCCACCTTGGGAGACGGGTCTTC 900
DB 2397 TGCTTGTCCCGGAGCCCCAGGGGCTCTCCCGGCTGCCACCTTGGGAGACGGGTCTTC 2456

QY 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960
DB 2457 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 2516

QY 961 CACCAACAGGAGGACCCAG 979
DB 2517 CACCAACAGGAGGACCCAG 2535
```

```
RESULT 10
ABA96945
ID ABA96945 standard; cDNA; 4463 BP.
XX
AC ABA96945;
XX
DT 20-MAY-2002 (first entry)
XX
DE Human ceramide kinase hCERK1-encoding cDNA.
XX
KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;
KW neurological disease; inflammation; human immunodeficiency virus;
KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;
KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;
KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;
KW gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 124..1737
FT FT /*tag= a
FT FT /product= "Human ceramide kinase hCERK1"
XX
PN WO200196575-A1.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-JP004889.
XX
PR 14-JUN-2000; 2000JP-00178039.
XX
(SANY ) SANKYO CO LTD.
XX
PI Sugiura M, Kono K, Kohama T;
XX
DR WPI; 2002-179513/23.
DR P-PSDB; AAM49115.
XX
PT Human ceramide kinase gene and the enzyme encoded by it for screening
XX substances as drugs for neurological, inflammatory and other disorders.
XX
PS Claim 5; Page 46-53; 61pp; Japanese.
XX
This sequence represents cDNA encoding a human ceramide kinase designated
hCERK1. The invention relates to hCERK1, nucleic acids encoding it,
expression vectors and host cells containing hCERK1 nucleic acids, the
recombinant production of hCERK1 and antibodies specific for hCERK1. The
invention also encompasses methods of isolating hCERK1 from samples, the
use of hCERK1 in drug screening, and the use of hCERK1 nucleic acid
sequences in gene therapy. hCERK1 mediates the ATP-dependent 1-
phosphorylation of ceramides and can be used to screen for therapeutic
and preventive agents for a wide range of disorders. Such disorders
include neurological disease, inflammation, human immunodeficiency virus
(HIV) infection, type 2 diabetes, obesity, sepsis, arteriosclerosis and
cancer
XX
SQ Sequence 4463 BP; 1026 A; 1135 C; 1217 G; 1084 T; 0 U; 1 Other;

Query Match      91.7%; Score 898.2; DB 6; Length 4463;
Best Local Similarity 96.3%; Pred. No. 6.8e-257;
Matches 943; Conservative 0; Mismatches 3; Indels 33; Gaps 1;

QY 1 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAAGACAAAGCAAGCGGAT 60
DB 510 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAAGACAAAGCAAGCGGAT 569

QY 61 ATATGAAAGAAAGTGGCACCACCTGTTTACCTTAGCTTCCATCACCCTGACATCATCGG 120
DB 570 ATATGAAAGAAAGTGGCACCACCTGTTTACCTTAGCTTCCATCACCCTGACATCATCG- 628

QY 121 TAACAAATTCATGTTAACTAGTAGAAGTAATTAAGCAATGCTAATCAGGCAAGGA 180
|||||
```


Db 629 -----TTACTGAACATGCTATATCAGGCCAAGA 656
Qy 181 GACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGG 240
Db 657 GACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGG 716
Qy 241 TATGTTTACGAGGTCGTCGACGGTCTGATTTGGGAGGACGACAGAGCGCGGGGTGCA 300
Db 717 TATGTTTACGAGGTCGTCGACGGTCTGATTTGGGAGGACGACAGAGCGCGGGGTGCA 776
Qy 301 CCAGAACACACCCCGGGTCTGCTGCTCCCAAGTAGCTCCGGATTGGAATCATTTCCCGC 360
Db 777 CCAGAACACACCCCGGGTCTGCTGCTCCCAAGTAGCTCCGGATTGGAATCATTTCCCGC 836
Qy 361 AGGGTCAACGGACTCGTGTGTTACTTCCACCGTGGGACCAAGCGACGAGAAACCTCGGC 420
Db 837 AGGTCCAAACGGACTCGTGTGTTACTTCCACCGTGGGACCAAGCGACGAGAAACCTCGGC 896
Qy 421 GCTGATATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCACAAACAG 480
Db 897 GCTGATATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCACAAACAG 956
Qy 481 CACACTCTCTTCTACTCTCGTCTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAA 540
Db 957 CACACTCTCTTCTACTCTCGTCTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAA 1016
Qy 541 GGACAGTGCAGAAAGACCGTGTGGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGAC 600
Db 1017 GGACAGTGCAGAAAGACCGTGTGGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGAC 1076
Qy 601 TTCTCTCTCCCAACCACTGCTATGAAAGGACAGTGTCTTCTCTCTGCAACAACACCGGT 660
Db 1077 TTCTCTCTCCCAACCACTGCTATGAAAGGACAGTGTCTTCTCTCTGCAACAACACCGGT 1136
Qy 661 GGGATCTCCAGGGATAGAAAGCCCTGCGGCGAGATGCTTTGTCGAGCAGCAAGCAA 720
Db 1137 GGGATCTCCAGGGATAGAAAGCCCTGCGGCGAGATGCTTTGTCGAGCAGCAAGCAA 1196
Qy 721 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAGCTGCGGAGGAGCT 780
Db 1197 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAGCTGCGGAGGAGCT 1256
Qy 781 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGCGCATCAATGCCAACAACATGCTCTG 840
Db 1257 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGCGCATCAATGCCAACAACATGCTCTG 1316
Qy 841 TGCTTTGTCGCGGAGCCCGAGGGCTCTCCCGGCTGCCACTTTGGGAGCGGCTCTTC 900
Db 1317 TGCTTTGTCGCGGAGCCCGAGGGCTCTCCCGGCTGCCACTTTGGGAGCGGCTCTTC 1376
Qy 901 TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 960
Db 1377 TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 1436
Qy 961 CACCAACACGAGGACAG 979
Db 1437 CACCAACACGAGGACAG 1455

RESULT 11

ADN62844

ID ADN62844 standard; DNA; 1740 BP.

XX AC ADN62844;

XX XX

DT 01-JUL-2004 (first entry)

XX DE Human NOV9a DNA.

XX ds; gene; human; NOVX; metabolic disorder; diabetes; obesity;

KW infectious disease; anorexia; cancer; cancer-associated cachexia;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

KW immune disorder; haematopoietic disorder; dyslipidaemia;

KW metabolic syndrome X; wasting disorder.

XX Homo sapiens.

XX US2004038223-A1.

XX 26-FEB-2004.

XX 01-OCT-2002; 2002US-00262511.

XX 02-OCT-2001; 2001US-0326483P.

XX 05-OCT-2001; 2001US-0327435P.

XX 05-OCT-2001; 2001US-0327449P.

XX 09-OCT-2001; 2001US-0327917P.

XX 09-OCT-2001; 2001US-0328029P.

XX 09-OCT-2001; 2001US-0328044P.

XX 09-OCT-2001; 2001US-0328056P.

XX 12-OCT-2001; 2001US-0328849P.

XX 15-OCT-2001; 2001US-0329414P.

XX 17-OCT-2001; 2001US-0330142P.

XX 18-OCT-2001; 2001US-0330309P.

XX 22-OCT-2001; 2001US-0341058P.

XX 24-OCT-2001; 2001US-0339266P.

XX 24-OCT-2001; 2001US-0343629P.

XX 29-OCT-2001; 2001US-0349575P.

XX 01-NOV-2001; 2001US-0346357P.

XX 17-APR-2002; 2002US-0373260P.

XX 19-APR-2002; 2002US-0373815P.

XX 19-APR-2002; 2002US-0373817P.

XX 19-APR-2002; 2002US-0373826P.

XX 19-APR-2002; 2002US-0373884P.

XX 22-APR-2002; 2002US-0374977P.

XX 16-MAY-2002; 2002US-0381037P.

XX 16-MAY-2002; 2002US-0381038P.

XX 16-MAY-2002; 2002US-0381042P.

XX 17-MAY-2002; 2002US-0381642P.

XX 28-MAY-2002; 2002US-0383656P.

XX 29-MAY-2002; 2002US-0383831P.

XX 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.

XX (MILL/) MILLET I.

XX (PEYM/) PEYMAN J A.

XX (KEKU/) KEKUDA R.

XX (JUJ/) JU J.

XX (LILL/) LI L.

XX (GUOX/) GUO X.

XX (PATT/) PATTURAJAN M.

XX (SPYT/) SPYTEK K A.

XX (EDIN/) EDINGER S R.

XX (ELLE/) ELLERMAN K.

XX (MALY/) MALYANKAR U M.

XX (ORTT/) ORT T.

XX (GORM/) GORMAN L.

XX (ZERH/) ZERHUSEN B D.

XX (ANDE/) ANDERSON D W.

XX (ZHON/) ZHONG M.

XX (CATT/) CATTERTON E.

XX (JIW/) JI W.

XX (MILL/) MILLER C E.

XX (RAST/) RASTELLI L.

XX (STON/) STONE D J.

XX (PENA/) PENNA C E A.

XX (SHEN/) SHENOY S G.

XX (SHIM/) SHIMKETS R A.

XX (ROTH/) ROTHENBERG M E.

XX (LEAC/) LEACH M D.

XX (AGEE/) AGEE M L.

XX (BERG/) BERGHS C.

XX (DIP/) DIPIPO V A.

XX (EISE/) EISEN A.

XX (GANG/) GANGOLLI E A.

XX (RIBG/) RIEGER D K.

(SPAD/) SPADERNA S K.

Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
Pacturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
Ott T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterson E;
Ji W, Miller CE, Rasceilli L, Stone DU, Pena CE, Shenoy SG;
Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
Eisen A, Gangolli EA, Rieger DK, Spaderna SK;
WPI: 2004-213931/20.
P-PSDB; ADN62845.

Isolated NOVX polypeptides and nucleic acids, useful for preventing,
diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

Claim 20; SEQ ID NO 39; 395pp; English.

The invention relates to isolated NOVX polypeptides and polynucleotides.
NOVX polypeptides and polynucleotides are used to prevent, diagnose or
treat a medical condition in human related to the aberrant expression and
activity of NOVX polypeptides. For example, NOVX polypeptides and
polynucleotides may be used to treat disorders associated with decreased
expression or activity of NOVX by supplementing the patient our
production or to rectify mutations. Conversely, antisense NA molecules
may be administered to down regulate expression of NOVX polypeptides by
binding with the cells own genes and preventing their expression. NOVX
polynucleotides and complementary sequences may also be used as DNA
probes in diagnostic assays to detect and quantitate the presence of
similar sequences in samples, and so which patients may be in need of
restorative therapy. NOVX polypeptides may also be used as antigens in
the production of antibodies and in assays to identify modulators
(agonists and antagonists) of the expression and activity of NOVX. The
anti-NOVX polypeptide antibodies, agonists and antagonists may also be
used to modulate NOVX polynucleotide expression and activity of NOVX
polypeptides. The anti-NOVX polypeptide antibodies may also be used as
diagnostic agents for detecting the presence of NOVX in samples. NOVX
polypeptides and polynucleotides may be used in this way to prevent,
diagnose and treat: metabolic disorders, diabetes, obesity, infectious
disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
haematopoietic disorders, and the various dyslipidaemias, metabolic
disturbances associated with obesity, the metabolic syndrome X and
wasting disorders associated with chronic diseases and various cancers.
They may also be used as antibacterial agents. The present sequence
represents DNA encoding a human NOVX protein.

Sequence 1740 BP; 388 A; 477 C; 530 G; 345 T; 0 U; 0 Other;

Query Match 89.7%; Score 878.4; DB 12; Length 1740;
Best Local Similarity 96.3%; Pred. No. 3.4e-251;
Matches 944; Conservative 0; Mismatches 1; Indels 35; Gaps 3;
1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGACAAGCAAGCGGAT 60
462 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGACAAGCAAGCGGAT 521
61 ATATGAAAGAAAGTGGACCACTGTTTCACTTACCTTACCTTCCATCCACTGACATCATCGG 120
522 ATATGAAAGAAAGTGGACCACTGTTTCACTTACCTTCCATCCACTGACATCATCG- 580
121 TAACAAATTTCTATTGTAACTATGTAGAAGTAATTAAGCAATGCTAATCAGGCAAGGA 180
581 -----TTACTGAACATGCTAATCAGGCAAGGA 608
181 GACTCTGTATGAGATTAAATAGACAATAATAGCGGCAATCGTCTGTTCGGCGGAGAT-G 239
609 GACTCTGTATGAGATTAAATAGACAATAATAGCGGCAAT-GTCTGTTCGGCGGAGATCG 667
240 GTATGTTTCAGCGAGTGTGCACCGTCTGATTGGGAGCAGCAGAGGAGCCCGGGTCG 299
668 GTATGTTTCAGCGAGTGTGCACCGTCTGATTGGGAGCAGCAGAGGAGCCCGGGTCG 727
300 ACCAAGAACCCCGGGCTGTGTGTCCTCCAGTAGCCTCCGGATTGGAATCATTCCTCG 359

Db 728 ACCAAGAACCCCGGGCTGTGTGTCCTCCAGTAGCCTCCGGATTGGAATCATTCCTCG 787
Qy 360 CAGGCTCAACCGAGTGGTGTGTACTTCCACCGTGGGCACAGCGAGCAAAACCTTCGG 419
Db 788 CAGGCTCAACCGAGTGGTGTGTACTTCCACCGTGGGCACAGCGAGCAAAACCTTCGG 847
Qy 420 CGCTGCATATCGTTGTTGGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACA 479
Db 848 CGCTGCATATCGTTGTTGGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACA 907
Qy 480 GCACATCTCTTCCGTACTCTCGTGTCTCTGCTGGGCTACGGCTTCTACGGGACATCATCA 539
Db 908 GCACATCTCTTCCGTACTCTCGTGTCTCTGCTGGGCTACGGCTTCTACGGGACATCATCA 967
Qy 540 AGGACAGTGAAGAAACCGTGTGGTCTGCCAGATACGACTTTTCAGGTTTAAAGA 599
Db 968 AGGACAGTGAAGAAACCGTGTGGTCTGCCAGATACGACTTTTCAGGTTTAAAGA 1027
Qy 600 CTTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCTCCCTGCACAAACACACGG 659
Db 1028 CTTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCTCCCTGCACAAACACACGG 1087
Qy 660 TGGGATCTCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTTTCAGGCAAAAGCA 719
Db 1088 TGGGATCTCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTTTCAGGCAAAAGCA 1147
Qy 720 AGCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAAAGCTTTCAGGATGCTCT 779
Db 1148 AGCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAAAGCTTTCAGGATGCTCT 1207
Qy 780 TGGAGGAGTGGCAAGTCTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCT 839
Db 1208 TGGAGGAGTGGCAAGTCTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCT 1267
Qy 840 GTGTTTGTGCGCGAGCCCCCAGGGGCTCTTCCCGGCTGCCACTTGGGAGACGGGTCTT 899
Db 1268 GTGTTTGTGCGCGAGCCCCCAGGGGCTCTTCCCGGCTGCCACTTGGGAGACGGGTCTT 1327
Qy 900 CTGACCTCATCTCCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGC 959
Db 1328 CTGACCTCATCTCCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGC 1387
Qy 960 ACACCAACCGAGGACAG 979
Db 1388 ACACCAACCGAGGACAG 1407
RESULT 12
AAS77728
ID AAS77728 standard; cDNA; 2241 BP.
XX AAS77728;
AC AAS77728;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13532.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
FA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-PSDB; ABG13541.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 13532; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2241 BP; 453 A; 627 C; 700 G; 461 T; 0 U; 0 Other;

Query Match 63.6%; Score 622.2; DB 5; Length 2241;
Best Local Similarity 81.6%; Pred. No. 1.3e-174;
Matches 827; Conservative 0; Mismatches 3; Indels 183; Gaps 2;

Qy 150 TAATTACTGAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAACTAGACAAAT 209
Db 896 TGATTACTGAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAACTAGACAAAT 955
Qy 210 ACGAGCGCATCTGTCTGTCGCGGAGATGTTATGTTCCAGCGAGTGTGCACCGTCTGA 269
Db 956 ACGAGCGCATCTGTCTGTCGCGGAGATGTTATGTTCCAGCGAGTGTGCACCGTCTGA 1015
Qy 270 TTGGGAGGACGACGAGGAGCGCGGGGTCGACCAAGAACCCCGCGGCTGTCTGGTCC 329
Db 1016 TTGGGAGGACGACGAGGAGCGCGGGGTCGACCAAGAACCCCGCGGCTGTCTGGTCC 1075
Qy 330 CCAGTAGCTCCGATTGGAATCATCTCCGCA----- 361
Db 1076 CCAGTAGCTCCGATTGGAATCATCTCCGCGAGAGCATGTGGGCCCGCGGTGAGAACCGTG 1135
Qy 362 -----GGGTCAACGAGCTCGCTGTGTACTCCACCGTGGGACGAGCGAG 410
Db 1136 GTGGCTTGACGGGTCAACGAGCTCGCTGTGTACTCCACCGTGGGACGAGCGAG 1195
Qy 411 AAACCTCGCGCTGCATATCGTTGTT----- 436
Db 1196 AAACCTCGCGCTGCATATCGTTGTTGGCTGTGTCGCCGAGCTCGGAACACCGCGCAT 1255
Qy 437 ----- 436
Db 1256 CCGGCCATACTGCTGTGGCAGTGGGCACGGGAGCTGTGCTGGCTGCCAGCTGTGA 1315
Qy 437 -----GGGACTCGC 446
Db 1316 ACGCATGCTCTGTGAGGCTCGAGGCTTCAGTCCAGGATCAGAGCCCCGGGACTCGC 1375

Qy 447 TGGCCATGGATGTGTCTCTCAGTCCACCAACAGCACATCTCTTGGTACTCTCGTGTCCC 506
Db 1376 TGGCCATGGATGTGTCTCTCAGTCCACCAACAGCACATCTCTTGGTACTCTCGTGTCCC 1435
Qy 507 TGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGGTGG 566
Db 1436 TGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGGTGG 1495
Qy 567 GTCTTGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCATCTGTATGAAG 626
Db 1496 GTCTTGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCATCTGTATGAAG 1555
Qy 627 GGACAGTGTCTCTCTCTCTCTGCAACACACAGGTGGGATCTCCAGGATAGGAGCCCT 686
Db 1556 GGACAGTGTCTCTCTCTCTCTGCAACACACAGGTGGGATCTCCAGGATAGGAGCCCT 1615
Qy 687 GCGGGGACGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGCAGAGA 746
Db 1616 GCGGGGACGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGCAGAGA 1675
Qy 747 AAGCACTGTATGTTTTCAGGAGCTGCGGAGGACGTTGGAGGAGTGGCAAGTCTGTGGGA 806
Db 1676 AAGCACTGTATGTTTTCAGGAGCTGCGGAGGACGTTGGAGGAGTGGCAAGTCTGTGGGA 1735
Qy 807 AGTTTCTGGCCATCAATGCCACAAACATGCTCTGCTGTCTGCGCGGAGCCCCAGGGGCC 866
Db 1736 AGTTTCTGGCCATCAATGCCACAAACATGCTCTGCTGTCTGCGCGGAGCCCCAGGGGCC 1795
Qy 867 TCTCCCGGCTGCCACCTTGGGAGCGGTCTTCTGACCTCATCTCATCGGAAATGCT 926
Db 1796 TCTCCCGGCTGCCACCTTGGGAGCGGTCTTCTGACCTCATCTCATCGGAAATGCT 1855
Qy 927 CCAGGTTCAATTTTCTGAGATTCTTCATCAGGCAACCAACAGCAGGACCCAG 979
Db 1856 CCAGGTTCAATTTTCTGAGATTCTTCATCAGGCAACCAACAGCAGGACCCAG 1908

RESULT 13
ADSI0370
ID ADSI0370 standard; DNA; 4702 BP.
XX
AC ADSI0370;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human therapeutic DNA - SEQ ID 607.
XX
XX antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
XX
OS Homo sapiens.
XX
XX WO2004080148-A2.
XX
PD 23-SEP-2004.
XX
PF 30-SEP-2003; 2003WO-US030720.
XX
PR 02-OCT-2002; 2002US-0416186P.
XX
PA (NUVE-) NUVELO INC.
XX
PI Tang YT, Asundi V, Ren F, Zhang J, Zhang J, Wehrman T, Wang Z, Ma Y;
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX
XX WPI; 2004-669857/65.
DR P-PSDB; ADSI11054.
XX
XX New polynucleotide, useful in preparing a composition for diagnosing or
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT aplastic anemia or cancer for promoting wound healing.
XX

PS Claim 1; SEQ ID NO 607; 718pp; English.

XX The invention relates to a novel isolated polynucleotide and the encoded
XX polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianemic, cytostatic and vulnerary activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during gene therapy
CC procedures. The current sequence is that of a human therapeutic DNA of
CC the invention. The current sequence is not shown explicitly within the
XX specification but can be accessed from the WIPO web-site.

SQ Sequence 4702 BP; 970 A; 1226 C; 1364 G; 1118 T; 0 U; 24 Other;

Query Match 63.1%; Score 618; DB 13; Length 4702;
Best Local Similarity 100.0%; Pred. No. 3.4e-173;
Matches 618; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 362 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACACGACGACGAGAACTCGGCG 421
Db 2044 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACACGACGACGAGAACTCGGCG 2103

Qy 422 CTGCATATCGTGTGGGACTCGCTGGCCATGATGTCCTCAGTCCACCAACAGC 481
Db 2104 CTGCATATCGTGTGGGACTCGCTGGCCATGATGTCCTCAGTCCACCAACAGC 2163

Qy 482 ACACCTCTTCTGCTACTCGGTGTCTCTGGCTACGGCTTCTACGGGGACATCATCAAG 541
Db 2164 ACACCTCTTCTGCTACTCGGTGTCTCTGGCTACGGCTTCTACGGGGACATCATCAAG 2223

Qy 542 GACAGTGAAGAAACGGTGTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 601
Db 2224 GACAGTGAAGAAACGGTGTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 2283

Qy 602 TTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACGCGTG 661
Db 2284 TTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACGCGTG 2343

Qy 662 GGATCTCCAGGGATAGGAGCCCTCGGGCAGGATGCTTTGTTTCAGGCAAGCAAG 721
Db 2344 GGATCTCCAGGGATAGGAGCCCTCGGGCAGGATGCTTTGTTTCAGGCAAGCAAG 2403

Qy 722 CAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGTG 781
Db 2404 CAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGTG 2463

Qy 782 GAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGSCCATCAATGCCAACAATGTCTCTGT 841
Db 2464 GAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGSCCATCAATGCCAACAATGTCTCTGT 2523

Qy 842 GCTTGTCCGGAGCCCGGAGGCTCTCCCGGCTGCCACTGGGAGACGGTCTTCT 901
Db 2524 GCTTGTCCGGAGCCCGGAGGCTCTCCCGGCTGCCACTGGGAGACGGTCTTCT 2583

Qy 902 GACCTCATCTCATCCGGAATGTCCAGGTTCAATTTTCTGAGATTTCATCAGGCAC 961
Db 2584 GACCTCATCTCATCCGGAATGTCCAGGTTCAATTTTCTGAGATTTCATCAGGCAC 2643

Qy 962 ACCAACGACGAGGACCCAG 979
Db 2644 ACCAACGACGAGGACCCAG 2661

RESULT 14
AAC76031
ID AAC76031 standard; cDNA; 753 BP.
XX
AC AAC76031;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1586 polynucleotide sequence SEQ ID NO:3171.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.

XX Homo sapiens.
OS
XX
XX WO200058473-A2.
PN
XX
XX 05-OCT-2000.
PD
XX
XX 31-MAR-2000; 2000WO-US008621.
PF
XX
XX 31-MAR-1999; 99US-0127607P.
PR
XX 02-APR-1999; 99US-0127636P.
PR
XX 05-APR-1999; 99US-0127728P.
PR
XX 30-MAR-2000; 2000US-00540763.
PR
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shimkets RA, Leach M;
PI
XX
XX WPI; 2000-602362/57.
DR
XX
XX P-PSDB; AAB41822.
DR

Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.

XX Claim 5; Page 2392; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX Sequence 753 BP; 165 A; 195 C; 207 G; 186 T; 0 U; 0 Other;

Query Match 61.2%; Score 598.8; DB 3; Length 753;
Best Local Similarity 99.7%; Pred. No. 7e-168;
Matches 600; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 360 CAGGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACACGACGAGAACTCGG 419
Db 131 CAGGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACACGACGAGAACTCGG 190

QY 420 CGCTGATATCGTTCTTGGGAGCTCGCTGGCCATGGATGTGTCTCAGTCACCAACA 479
Db |||||
191 CGCTGATATCGTTGTGGGAGCTCGCTGGCCATGGATGTGTCTCAGTCACCAACA 250
QY 480 GCACACTCTTCGCTACTCCGTGTCCCTGCTGGGCTACCGGCTTCTACGGGACATCATCA 539
Db |||||
251 GCACACTCTTCGCTACTCCGTGTCCCTGCTGGGCTACCGGCTTCTACGGGACATCATCA 310
QY 540 AGGACAGTGAAGAAACCGTGGTGGTCTTTGCCAGATACGACTTTTCAGGTTTAAAGA 599
Db |||||
311 AGGACAGTGAAGAAACCGTGGTGGTCTTTGCCAGATACGACTTTTCAGGTTTAAAGA 370
QY 600 CTTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACGG 659
Db |||||
371 CTTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACGG 430
QY 660 TGGGATCTCCAGGATAGGAAGCCCTCGCGGCAGGATGCTTTGTTGCGAGCAAGCA 719
Db |||||
431 TGGGATCTCCAGGATAGGAAGCCCTCGCGGCAGGATGCTTTGTTGCGAGCAAGCA 490
QY 720 AGCAGAGCTGAGGAGGACGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACG 779
Db |||||
491 AGCAGAGCTGAGGAGGACGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGATG 550
QY 780 TGGAGAGTGGCAAGTCTGTGGGAAGTTTCTGCCATCAATGCCACAAACATGTCT 839
Db |||||
551 TGGAGAGTGGCAAGTCTGTGGGAAGTTTCTGCCATCAATGCCACAAACATGTCT 610
QY 840 GTGCTTGTCCGCGAGCCCGAGGGCTCTCCCGGCTGCCACATTTGGAGACGGTCT 899
Db |||||
611 GTGCTTGTCCGCGAGCCCGAGGGCTCTCCCGGCTGCCACATTTGGAGACGGTCT 670
QY 900 CTGACCTCATCTCATCCGGAATGCTCCAGTTTCAATTTCTGAGATTTCTCATCAGGC 959
Db |||||
671 CTGACCTCATCTCATCCGGAATGCTCCAGTTTCAATTTCTGAGATTTCTCATCTGC 730
QY 960 AC 961
Db ||
731 AC 732
RESULT 15
ID AAS77731 standard; cDNA; 1570 BP.
AC AAS77731;
XX AAS77731;
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13535.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
PF
XX
XX 31-MAR-2000; 2000US-00540217.
PR
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI; 2001-639362/73.
DR
XX P-PSDB; ABG13544.
DR
XX
XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 13535; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful for medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1570 BP; 349 A; 446 C; 456 G; 319 T; 0 U; 0 Other;
Query Match 55.5%; Score 543; DB 5; Length 1570;
Best Local Similarity 100.0%; Pred. No. 4.8e-151;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 437 GGGGACTCGCTGGCCATGGATGTCTCTCAGTCACCAACAGCAGTCTCTCGCTAC 496
Db |||||
595 GGGGACTCGCTGGCCATGGATGTCTCTCAGTCACCAACAGCAGTCTCTCGCTAC 754
QY 497 TCCGTGTCTCTGTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAA 556
Db |||||
755 TCCGTGTCTCTGTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAA 814
QY 557 CGGTGGTGGTCTTGGCAGATACGCTTTTTCAGGTTTAAAGACTTCTCTCCACAC 616
Db |||||
815 CGGTGGTGGTCTTGGCAGATACGCTTTTTCAGGTTTAAAGACTTCTCTCCACAC 874
QY 617 TGCTATGAAGGACAGTGTCTCTCTCCGCAACACACGCTGGGATCTCCAAGGAT 676
Db |||||
875 TGCTATGAAGGACAGTGTCTCTCTCCGCAACACACGCTGGGATCTCCAAGGAT 934
QY 677 AGGAAGCCCTGCCGGGAGGATGCTTTGTTTGCAGGCAAGCAAGCAGCTGGAGGAG 736
Db |||||
935 AGGAAGCCCTGCCGGGAGGATGCTTTGTTTGCAGGCAAGCAAGCAGCTGGAGGAG 994
QY 737 GAGCAGAAAGAACACTGTATGTTTGAAGCTCGGAGGACGCTGGAGGATGGCAATC 796
Db |||||
995 GAGCAGAAAGAACACTGTATGTTTGAAGCTCGGAGGACGCTGGAGGATGGCAATC 1054
QY 797 GTCTGTGGAGTGTCTTGGCCATCAATGCCAACACATGTCTGTGTGTGCGCGGAGC 856
Db |||||
1055 GTCTGTGGAGTGTCTTGGCCATCAATGCCAACACATGTCTGTGTGTGCGCGGAGC 1114
QY 857 CCCAGGGGCTCTCTCCCGGCTGCCACATTTGGGAGACGGGTCTTCTGACCTCATCTCATC 916
Db |||||
1115 CCCAGGGGCTCTCTCCCGGCTGCCACATTTGGGAGACGGGTCTTCTGACCTCATCTCATC 1174
QY 917 CGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATAGGACACCAACAGGAGAC 976
Db |||||
1175 CGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATAGGACACCAACAGGAGAC 1234
QY 977 CAG 979
Db ||||
1235 CAG 1237

Search completed: September 5, 2005, 20:25:07
Job time : 578.067 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 09:12:32 ; Search time 179.716 Seconds
(without alignments)
8913.583 Million cell updates/sec

Title: US-10-631-958-1

Perfect score: 979

Sequence: 1 accaaagcattactgtgtat.....acaccaaccagcaggaccag 979

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|----|---------------------|
| 1 | 901.4 | 92.1 | 4432 | 4 | US-09-774-528-148 |
| 2 | 78.6 | 8.0 | 1084 | 4 | US-09-270-767-15155 |
| 3 | 77 | 7.9 | 2064 | 4 | US-09-270-767-14306 |
| 4 | 69.8 | 7.1 | 498 | 4 | US-09-893-737-59 |
| 5 | 55 | 5.6 | 901 | 4 | US-09-270-767-30448 |
| 6 | 43.4 | 4.4 | 1050 | 4 | US-09-252-991A-1269 |
| 7 | 43.4 | 4.4 | 1308 | 4 | US-09-252-991A-1179 |
| 8 | 42.2 | 4.3 | 430 | 4 | US-09-621-976-16656 |
| 9 | 40.8 | 4.2 | 1014 | 4 | US-09-902-540-9500 |
| 10 | 40.8 | 4.2 | 14101 | 4 | US-09-902-540-1080 |
| 11 | 40.4 | 4.1 | 1857 | 4 | US-09-970-516-3 |
| 12 | 40.4 | 4.1 | 2380 | 4 | US-09-817-676A-13 |
| 13 | 39.8 | 4.1 | 485 | 4 | US-09-270-767-31476 |
| 14 | 39.4 | 4.0 | 2698 | 4 | US-09-817-676A-11 |
| 15 | 38.8 | 4.0 | 1875 | 4 | US-09-614-221A-399 |
| 16 | 38.4 | 3.9 | 2274 | 4 | US-09-252-991A-3660 |
| 17 | 38.4 | 3.9 | 3147 | 4 | US-09-902-540-3530 |
| 18 | 38.4 | 3.9 | 3297 | 4 | US-09-252-991A-3615 |
| 19 | 38.4 | 3.9 | 18538 | 4 | US-09-902-540-1169 |
| 20 | 37.8 | 3.9 | 395 | 4 | US-09-894-844-45 |
| 21 | 37.8 | 3.9 | 4403765 | 3 | US-09-103-840A-2 |
| 22 | 37.8 | 3.9 | 4411529 | 3 | US-09-103-840A-1 |
| 23 | 37.2 | 3.8 | 6644 | 4 | US-08-875-435B-5 |
| 24 | 36.8 | 3.8 | 601 | 4 | US-09-949-016-62772 |
| 25 | 36.8 | 3.8 | 601 | 4 | US-09-949-016-62773 |
| 26 | 36.8 | 3.8 | 103792 | 4 | US-09-949-016-13553 |
| 27 | 36.8 | 3.8 | 103993 | 4 | US-09-949-016-12317 |

| | | | | | | |
|------|------|-----|--------|---|---------------------|-------------------|
| 28 | 36.6 | 3.7 | 601 | 4 | US-09-949-016-53419 | Sequence 53419, A |
| c 29 | 36.6 | 3.7 | 139936 | 4 | US-09-949-016-11782 | Sequence 11782, A |
| c 30 | 36.6 | 3.7 | 139952 | 4 | US-09-949-016-13280 | Sequence 13280, A |
| 31 | 36.4 | 3.7 | 1086 | 4 | US-08-489-039A-1943 | Sequence 1943, Ap |
| c 32 | 36.4 | 3.7 | 7218 | 1 | US-08-232-463-14 | Sequence 14, Appl |
| 33 | 36 | 3.7 | 3163 | 4 | US-09-949-016-3948 | Sequence 3948, Ap |
| 34 | 36 | 3.7 | 3171 | 4 | US-09-949-016-202 | Sequence 202, App |
| 35 | 36 | 3.7 | 44653 | 4 | US-09-949-016-11944 | Sequence 11944, A |
| 36 | 36 | 3.7 | 44653 | 4 | US-09-949-016-15690 | Sequence 15690, A |
| 37 | 35.8 | 3.7 | 1588 | 4 | US-09-490-291-7 | Sequence 7, Appli |
| 38 | 35.8 | 3.7 | 3709 | 3 | US-09-541-782-7 | Sequence 7, Appli |
| 39 | 35.8 | 3.7 | 3709 | 3 | US-09-723-820-7 | Sequence 7, Appli |
| 40 | 35.8 | 3.7 | 3709 | 4 | US-10-270-085-7 | Sequence 7500, Ap |
| c 41 | 35.6 | 3.6 | 1821 | 4 | US-09-902-540-7500 | Sequence 1, Appli |
| 42 | 35.6 | 3.6 | 3946 | 3 | US-09-083-351-1 | Sequence 1, Appli |
| 43 | 35.6 | 3.6 | 3946 | 3 | US-09-083-352-1 | Sequence 1, Appli |
| 44 | 35.6 | 3.6 | 5040 | 4 | US-09-902-540-727 | Sequence 727, App |
| 45 | 35.4 | 3.6 | 3468 | 1 | US-07-951-715A-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-09-774-528-148
; Sequence 148, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 148
; LENGTH: 4432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129)..(2817)
US-09-774-528-148

| | | | | | | | |
|-----------------------|-------|-----------------------------------------------------------|----------|------------|---|--------|------|
| Query Match | 92.1% | Score | 901.4 | DB | 4 | Length | 4432 |
| Best Local Similarity | 96.5% | Pred. No. | 1.4e-266 | | | | |
| Matches | 945 | Conservative | 0 | Mismatches | 1 | Indels | 33 |
| Gaps | 1 | | | | | | |
| Qy | 1 | ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACAAAGCGGAT | 60 | | | | |
| Db | 1590 | ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACAAAGCGGAT | 1649 | | | | |
| Qy | 61 | ATATGAAGAAAGTGGACCACTGTTACCTTAGCCTCCATCACCCTGATCATCGG | 120 | | | | |
| Db | 1650 | ATATGAAGAAAGTGGACCACTGTTACCTTAGCCTCCATCACCCTGATCATCGG | 1708 | | | | |
| Qy | 121 | TACAAATTCCTGTTAACTAGTAAGTAATTAATGAACTGTAATCAATCAAGGCAAGGA | 180 | | | | |
| Db | 1709 | TTACTGAACATGCTAATCAAGGCAAGGA | 1736 | | | | |

| | | | |
|-----------------------------------------------------------------------------|-----|-----------------------------------------------------------------|-----|
| | | ; ORGANISM: Drosophila melanogaster | |
| | | US-09-270-767-15155 | |
| | | Query Match 8.0%; Score 78.6; DB 4; Length 1084; | |
| | | Best Local Similarity 49.1%; Pred. No. 1.5e-13; | |
| | | Matches 238; Conservative 0; Mismatches 244; Indels 3; Gaps 1; | |
| Qy | 150 | TAATTACTGAACATGCTAAATCAGGCCAAAGGAGACTCTGTATGAGATTAAACATAGACAAAT | 209 |
| Db | 926 | TCACCACTCAGAGGGCAACCAAGTGAAGACATACCTCTTGAGCCATGATCTGGAGTAT | 867 |
| Qy | 210 | ACGAGCGCATCGTCTGTGTCGGCGGAGATGGTATGTTTCAGCGAGGTCTGCACGGTCTGA | 269 |
| Db | 866 | ACGATGCGGTTTGTGTGTCGGAGCGATGCCACCGTAGCAGAGGTCAACAACGACTGA | 807 |
| Qy | 270 | TTGGGAGGACGACAGAGAGCGCGGGGTGCACCAAGAACCCCGGGCTGTGTGTCTCC | 329 |
| Db | 806 | TAATTCCTCAAAATCCGAGAGTTGGGACTGGACGAACAGCGGCCACCATACATTCCAAAGAC | 747 |
| Qy | 330 | CCAGTAGCCCTCCGGATTGGAAATCATTCGCCGAGGGTCAACGGACTGGTGTGTACTCCA | 389 |
| Db | 746 | CGGT----CTGCCAGTGGGTGTGATTCGCCGTGGCAGCACCGACCATTTGCGTATAGTA | 690 |
| Qy | 390 | CCGTGGCCACACGACGACGAGAAACCTCGCGCTGCATATCGTTTGTGGGACTCGCTGG | 449 |
| Db | 689 | TGCACGGCACCGGGGATGTGAGGACAGCGCTATCCATGTGATTCTGGGCCAGCATCGG | 630 |
| Qy | 450 | CCATGATGTGCTCTAGTCCACCAACAGCACACTCTCTCGCTACTCCGTGTCTCCCTGC | 509 |
| Db | 629 | GATTGGATGTGTGCAGTGTGAGCAATGGCCAGTCCCTGCTCAGATTCTGTGCCAGTGTCC | 570 |
| Qy | 510 | TGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAGAGAAACGGTGGTGGGTG | 569 |
| Db | 569 | TGAGCTACGGTACTCTGGCGATGTGGCAGCCAGAGCGAGACTACCGCTGGATGGGAC | 510 |
| Qy | 570 | TTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGTCTATGAAGGA | 629 |
| Db | 509 | CGCGCCGTACGAGTACAGTGGCGTCAAGGCTTCTCCTGAATAATCGCGGCTATGACCGG | 450 |
| Qy | 630 | CAGTG 634 | |
| Db | 449 | AAC TG 445 | |
| RESULT 3 | | | |
| US-09-270-767-14306 | | | |
| ; Sequence 14306, Application US/09270767 | | | |
| ; Patent No. 6703491 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Homburger et al. | | | |
| ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster | | | |
| ; FILE REFERENCE: File Reference: 7326-094 | | | |
| ; CURRENT APPLICATION NUMBER: US/09/270,767 | | | |
| ; CURRENT FILING DATE: 1999-03-17 | | | |
| ; NUMBER OF SEQ ID NOS: 62517 | | | |
| ; SOFTWARE: PatentIn Ver. 2.0 | | | |
| ; SEQ ID NO 14306 | | | |
| ; LENGTH: 2064 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Drosophila melanogaster | | | |
| US-09-270-767-14306 | | | |
| | | Query Match 7.9%; Score 77; DB 4; Length 2064; | |
| | | Best Local Similarity 48.9%; Pred. No. 7.2e-13; | |
| | | Matches 237; Conservative 0; Mismatches 245; Indels 3; Gaps 1; | |
| Qy | 150 | TAATTACTGAACATGCTAAATCAGGCCAAAGGAGACTCTGTATGAGATTAAACATAGACAAAT | 209 |
| Db | 722 | TCACCACTCAGAGGGCAACCAAGTGAAGACATACCTCTTGAGCCATGATCTGGAGTAT | 781 |
| Qy | 210 | ACGAGCGCATCGTCTGTGTCGGCGGAGATGGTATGTTTCAGCGAGGTCTGCACGGTCTGA | 269 |
| Db | 782 | ACGATGCGGTTTGTCTGTGTCGGAGCGATGGCCCGTAGCAGAGGTCAACAACGACTGA | 841 |

QY 270 TTGGAGGACGAGAGGAGCGCGGGGTGACACGAAACACCCCGGGGTGTGTGGTCC 329
DB 842 TATTCGGTCAAAATGCGAGAGTTGGGACTGGACGAAACAGCGGCCACATACATTCCAAGAC 901
QY 330 CCAGTAGCTCCGGATTGGAATCATTTCCCGCAGGGTCAACGAGCTGCGTGTACTCCA 389
DB 902 CGGCT---CTGCAGTGGGTGTGATTCCTCCGCTGGCAGCACCGACACCATTCGGGTATAGTA 958
QY 390 CGGTGGGACACGACGACGAGAAACCTCGGCGCTGCATATCTGTTGGGGACTCGCTGG 449
DB 959 TGCACGCGACGCGGATGTGAGGACAGCGGCTATCCATGTGATTCGGGCGAGCATCGG 1018
QY 450 CCATGGATGTTCCTCAGTCCACCAACAGACACACTCTCTCGCTACTCCGTTCCCTGC 509
DB 1019 GATTGGATGTGTGCAGTGTGAGCAATGGCGAGTCCCTGCTCAGATTCTGTGCCAGTGTCC 1078
QY 510 TGGGCTACGGCTTCTACGCGGACATCATCAAGGACAGTGAGAAGAACGGTGTGGGTC 569
DB 1079 TGAGCTACGGTACCTGGCGGATGTGGAGCCGAGCGAGAACTACCGCTGGATGGGAC 1138
QY 570 TTGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGA 629
DB 1139 CGCGCGGTACGAGTACAGTGGCGTCAAGGCTTCTGTAATATCGCGGCTATGACGCG 1198
QY 630 CAGTG 634
DB 1199 AACTG 1203

RESULT 4

US-09-893-737-59/c
; Sequence 59, Application US/09893737
; Patent No. 6822082
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; PRIORITY FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(498)
US-09-893-737-59

Query Match 7.1%; Score 69.8; DB 4; Length 498;
Best Local Similarity 91.4%; Pred. No. 4.8e-11;
Matches 74; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 360 CAGGCTCAACGAGCTGCGTGTGTTACTCCACGTTGGGACCGACGACGAGAACTCGG 419
DB 460 CAGGCTCAACGAGCTGCGTGTGTTACTCCACGTTGGGACCGACGACGAGAACTCGG 401
QY 420 CGCTGCATATCGTTGTTGGG 440
DB 400 CGCTGCATATCGTTGTTGGG 380

RESULT 5

US-09-270-767-30448
; Sequence 30448, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30448
; LENGTH: 901
; TYPE: DNA
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-30448

Query Match 5.6%; Score 55; DB 4; Length 901;
Best Local Similarity 57.1%; Pred. No. 2.5e-06;
Matches 100; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 782 GAGCAGTGGCAAGTCGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTGT 841
DB 368 GATCAGTGGAAAGTTGTGGCGGCAATTTCTTTATGATCTCGGGCGCAATACCTGC 427
QY 842 GTTTGTGCGCGGAGCCCGGCGCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTCT 901
DB 428 GCCTGCGCAGGAGTCCCAATGGCATCTCCCGTTACAGTCATCTGGGTGATGTTGCCCTG 487
QY 902 GACCTATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCA 956
DB 488 GACCTGATTTCTCGTGAAGAACCTCACTGCTCAACACGTCGCTTTTCTGCTCA 542

RESULT 6

US-09-252-991A-1269
; Sequence 1269, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIORITY FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1269
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-1269

Query Match 4.4%; Score 43.4; DB 4; Length 1050;
Best Local Similarity 49.3%; Pred. No. 0.01;
Matches 113; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 307 CCACCCCCGGGCTGTGCTGGTCCCAGTAGCTCCGAGTTGGAATCATTTCCCGAGGGTC 366
DB 691 CCATCTATGTTGGGTCTACGCCACACAGTCTGCTCTGATCGGCGCACCTTTCGCGGCTT 750
QY 367 AACGGATGGTGTGTTACTCCACCGTGGGACAGGACGACGAGAACTCGGCGCTGCA 426
DB 751 CACCTATTTCTGTCGATCTCTCACCGAGGTGAGCGGTTCTCCGCTCGACCGTACCGCT 810
QY 427 TATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCACACACACACACACT 486
DB 811 GCTGCTGGTGTCTACCGGCTTGGGACGCTGGTGGGCAACATCTGCGGCCCTGGC 870
QY 487 CTTTCGCTACTCCGTGTCTCTGCTGGGCTACGGCTTCTACGGGACATC 535
DB 871 CGACCGCATACCATCGGGTCTCTGGCTTCGGGCTGCTGGCGCCATC 919

RESULT 7

```
US-09-252-991A-1179
; Sequence 1179, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1179
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1179

Query Match          4.4%; Score 43.4; DB 4; Length 1308;
Best Local Similarity 49.3%; Pred. No. 0.012;
Matches 113; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

Qy 307 CCACCCCGGCTGTGCTGGTCCCGAGTAGCCTCCGGATTGGAATCATTCGCCGAGGGTC 366
Db 741 CCACTATGTGGGTCTACGGACACAGTCTGCTGTATCGCGCCACCTTCGCCGGCTT 800

Qy 367 AACGGACTGCGTGTGTTACTCCACCGTGGGACCGAGCAGCAAGAACTTCGGCGCTGCA 426
Db 801 CACCTATTTCGTGGCGATCCTCACCGAGGTTCAGCGGCTTCCTCCGCTCGACCGTACCGCT 860

Qy 427 TATCGTGTGGGACTCGTGGCCATGGATGTCTCAGTCCAGTCCACCAACAGCACACT 486
Db 861 GCTGCTGTGTGTCTACGGCCCTGGGACGCTGGTGGGCAACAACTCGTCGCCCGCCTGGC 920

Qy 487 CTTTCGCTACTCCGTGTCCTCTGGGCTACGGCTTCTACGGGACATC 535
Db 921 CGACGCCATACCATCGCGGTCTGGCCCTTCGGCCCTGCTGGCGGCCATC 969

RESULT 8
US-09-621-976-16656
; Sequence 16656, Application US/09621976
; Patent No. 6839063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16656
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16656

Query Match          4.3%; Score 42.2; DB 4; Length 430;
Best Local Similarity 19.2%; Pred. No. 0.014;
Matches 50; Conservative 99; Mismatches 112; Indels 0; Gaps 0;

Qy 640 CTTCCCTGCACAACACAGCTGGGATCTCCAGGGATAGAGCCCTCCGGCGAGGATG 699
Db 35 MNCMMRSCCYMWCSSGRRSCSRAMCCYYKKKSCSRAMCCCYTYCYKSSCCSYKGS 94

Qy 700 CTTTCTTTGGAGGAAAGAACGACGACGCTGGAGGAGGAGCAAGAAAGCACTGTATGG 759
Db 95 YTTKRAMMWRKRSCTYSRRRRYYYRSMRSMWKGSCCCSCSGGSCYKXKKKKGKGC 154
```

```
Qy 760 TTTGGAAGCTGCGAGGACGTGAGGAGTGGCAAGTCTGTCTGTGGGAAGTTTCTGGCCAT 819
Db 155 CMRSYWMCCYYRARRRMWKGSGSCHMYTKRMRRMCCCCHRRRSMRRRMCMWKGSYTY 214

Qy 820 CAATGCCACAACAATGCTCTGTGTCGCCGAGAGCCCGAGGGGCTCTCCCGGGTGC 879
Db 215 CYKSSMMCMARRKARAGKRMCCYTKGGGRMMRYCCMRKGRRACTCTGTTCACTGC 274

Qy 880 CCACCTGGGAGACGGGTCTTC 900
Db 275 CGTGTCTGCCAGAAGGCGCTTC 295

RESULT 9
US-09-902-540-9500
; Sequence 9500, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9500
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-9500

Query Match          4.2%; Score 40.8; DB 4; Length 1014;
Best Local Similarity 46.2%; Pred. No. 0.064;
Matches 135; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

Qy 215 GGCATCGTCTGTGTCGGCGGAGATGGTATGTTACGCGAGGTGCTGCACGGTCTGATGGG 274
Db 679 GGCAGCCCGCTGCTGGGCAACCCCGTGATCTTCGCGAGCTGCTCGAGGCGCCCGCCGC 738

Qy 275 AGGACGACAGGAGCGCGGCTCGACCAACACCCCGGGCTGTGCTGGTCCCGCAGT 334
Db 739 AGCCCGCGAGCGCTCGAGCTGTGCTGAGACATTCGCGGCGACCTGGACTTCATG 798

Qy 335 AGCCTCCGGATTGGAATCATTCGCGAGGGTCAACGGAATGCGTGTGTTACTCCACCGTG 394
Db 799 GGGGACCGCTGGGCGCGCTGCGCTCTTCGCAAGCAGTTGGCGTGTGTACGCCACGGC 858

Qy 395 GGACACGAGCGACGAGAACTCGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATG 454
Db 859 TTGTACGCGCGCGCGCTTCGCTGCGAGGTGAACCGCTTGGACCTGCGCGTCAAGCGGTG 918

Qy 455 GATGTGCTCTCAGTCCACCAACACGACACACTCTCTCGCTACTCCGTGTCCC 506
Db 919 GAGGACTGCGTGGCGCGCTTCTTCGCGCGGCCCATGTGGACCTCGCGGCC 970

RESULT 10
US-09-902-540-1080/c
; Sequence 1080, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
```

```
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1080
; LENGTH: 14101
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1080

Query Match
Best Local Similarity 4.2%; Score 40.8; DB 4; Length 14101;
Matches 135; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 215 GGCATCGTCTGTGTCGGCGAGATGGTATGTTTCACGAGGTGCTCGACGGTCTGATGCGG 274
Db 9788 GCGCCGCGCGCTGCGCGAACCCTGGATCTTCGCGAGCTCTCGGAGGCGCGCCGCC 9729

QY 275 AGGACGCGAGAGCGCGGGTGCACAGAACACCCCGGCTGTGCTGGTCCCCAGT 334
Db 9728 AGCCGCGGAGCGCTGCGAGCTGTGTGGAGCACTTCGCGGCGCACCTGGACTTCATG 9669

QY 335 AGCCTCCGGAATGGAATCATTCGCCGAGGTCACCGACTGCGTGTGTTACTTCCACCGTG 394
Db 9668 GGGGACCCGCTGGGCGCGCTCGCTCTCCGCAAGCATGGCGTGGTACGCCACCGC 9609

QY 395 GGCACACGACGACGAGAACCTCGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATG 454
Db 9608 TTGTACGCGCGCGGCTTCGTCGCGAGGTGAACGCGCTGGACCTGCCGTGACGCGTG 9549

QY 455 GATGTGCTCTAGTCCACCAACAGCACACTCTTCCTGCTACTCGTGTCTCC 506
Db 9548 GAGGACTCGTCCGCGCTTCTTCGCGCGGCCCATGTGGACCTCGCGGGCC 9497

RESULT 11
US-09-970-516-3
; Sequence 3, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1857)
; OTHER INFORMATION:
US-09-970-516-3

Query Match
Best Local Similarity 4.1%; Score 40.4; DB 4; Length 1857;
Matches 219; Conservative 0; Mismatches 266; Indels 9; Gaps 1;

QY 155 ACTGAACATGCTAATCAGGCCAAGAGACTCTGTATGAGATTAAACATAGACAAATACGAC 214
Db 550 ACAGAACACAGAACACACCGGAGCTGTGTCAGGGGCTGAGCTGAGTGAAGTGGAT 609

QY 215 GGCATCGTCTGTGTCGGCGAGATGGTATGTTTCACGAGGTGCTGCACGGTCTGATGGG 274
Db 610 GGCATCGTCAAGTCTCGGCGAGACGCGGCTGTCTCCATGAGGTGCTGAACGGGCTCTAGAT 669

QY 275 AGGACGCGAGAGCGCGGGTGCACAGAACACCCCGGCTGTGCTGGTCCCCAGT 334
Db 670 GCGCCTGACTGGGAGAGCTGTGAAGATG-----CCTGTGGGCATCCTCCCCCTGC 720
```

```
QY 335 AGCCTCCGGAATGGAATCATTCGCCGAGGTCACCGACTGCGTGTGTTACTTCCACCGTG 394
Db 721 GGCCTCGGCAACACCGCTGCGCGAGCAGTGAACACGACACGCGGGGATTTGAGCAGCCCTG 780

QY 395 GGCACACGACGACGAGAACCTCGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATG 454
Db 781 GGCCTCGACCTGTTGCTCACTGTTGTCGCCGGGGTGGGGCCACCCACTG 840

QY 455 GATGTGCTCTCAGTCCACCAACAGCACACTCTTCCTACTCTCGTGTCCCTGCTGGGC 514
Db 841 GACCTGCTCTCCGTGACGCTCGGCTCGGCTCGGCTGTTCTCTCTCTCTCTCTGTCG 900

QY 515 TACGGCTCTACGGGGACATCATCAAGGACAGTGAAGAAACGGTGTGGTCTTCC 574
Db 901 TGGGCTCTCGTGTGATATCCAGAGCAGCGCTTCAGGGCTTGGGCAAGTGCC 960

QY 575 AGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACATG 634
Db 961 CGCTTACACTGGGACGCGTGTGGGCTCGCCACACTGACACCTACCGCGGACGCGCTC 1020

QY 635 TCCTTCTCTCCCTGC 648
Db 1021 TCCTACCTCCCGC 1034
```

RESULT 12

```
US-09-817-676A-13
; Sequence 13, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1860)
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245447
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-13
```

```
Query Match
Best Local Similarity 4.1%; Score 40.4; DB 4; Length 2380;
Matches 219; Conservative 0; Mismatches 266; Indels 9; Gaps 1;

QY 155 ACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGAC 214
Db 556 ACAGAACACAGAACACACCGGAGCTGTGTCAGGGGCTGAGCTGAGTGAAGTGGAT 615

QY 215 GGCATCGTCTGTGTCGGCGAGATGGTATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGG 274
Db 616 GGCATCGTCAAGTCTCGGAGAGCGGCTGTCTCATGAGGTGCTGAACGGGCTCTAGAT 675
```

```
Qy 275 AGGACGACAGAGGAGCGCGGGTCGACAGAACACCCCGGGCTGTCTGGTCCCCAGT 334
    |||||
Db 676 CGCCCTGACTGGGAGGAGCTGTGAATG-----CCTGTGGGCATCTCCCTGC 726

Qy 335 AGCTCCGGATTGGAATCATTTCCGCGAGGGTCAACGAGCTCGGTGTACTCCACCGTG 394
    |||||
Db 727 GGCTCGGGCAACGGCTGGCCGAGCAGTGAACAGCAGCGGGGATTTGAGCCAGCCCTG 786

Qy 395 GGCACACGACGACGAGAAACCTCGGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATG 454
    |||||
Db 787 GGCCTCGACCTGTTGCTCACTGTCTGCTGTGTCGCGGGTGGTGGCCACCCACTG 846

Qy 455 GATGTGCTCAGTCCACCAACAGCACACTCTTCCGCTACCTCGGTGTCCTGCTGGCC 514
    |||||
Db 847 GACCTGCTCTCCGTGACGCTGGCCTCGGGCTCCCGCTGTTCTCTCCTCTGTGTGCC 906

Qy 515 TACGGCTTCTACGGGGACATCATCAAGGACAGTGAAGAAACCGTGTGGTCTTGGCC 574
    |||||
Db 907 TGGGGCTTCGTGTGAGATGTGGATATCCAGAGCGAGCGCTTCAGGGCCCTTGGGCAGTGCC 966

Qy 575 AGATACGACTTTTACAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTG 634
    |||||
Db 967 CGCTTCACTGGGACGGTGTGGGCTCGCCACACTGACACCTACCGCGGAGCCCTC 1026

Qy 635 TCCTTCTCCCTGC 648
    |||||
Db 1027 TCCTACCTCCCGC 1040
```

RESULT 13

```
US-09-270-767-31476/c
; Sequence 31476, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; PRIORITY FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 31476
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-31476
```

```
Query Match 4.1%; Score 39.8; DB 4; Length 485;
Best Local Similarity 59.1%; Pred. No. 0.081;
Matches 68; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 782 GAGGAGTGGCAAGTCTGTGGGAGTTTCTGGCCATCATGCCACAAACATGTCTGT 841
Db 117 GATCAGTGGAGGTTGTGCGGGCAATTTCTTTATGATCTCGCGCGCAACATAACCTGC 58

Qy 842 GCTTGTCCCGAGGCGCCAGCGGCTCTCCCGGCTGCCACTTGGGAGACGGT 896
Db 57 GCCTGGCCAGAGTCCCAATGGCAATCTCCGGTTACAGTCACTCGGTGATGTTT 3
```

RESULT 14

```
US-09-817-676A-11
; Sequence 11, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: Expression and Methods of Use Thereof
; CURRENT APPLICATION NUMBER: US/09/817,676A
; PRIORITY FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
```

```
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2698
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (387)..(2237)
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245448
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-11
```

```
Query Match 4.0%; Score 39.4; DB 4; Length 2698;
Best Local Similarity 55.5%; Pred. No. 0.32;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 150 TAATTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAAT 209
    |||||
Db 934 TACAGACAGAACGACAGAACCATGCCCGTGGTGTGAGGGTTAAGCTGAGTGAGT 993

Qy 210 ACGACGGCATCGTGTGTGTCGGCGGAGATGTTTTCAGCGAGGTCTGCACGGTCTGA 269
    |||||
Db 994 GGGAGGSCATTGTTCACGTGTCTGGAGACGGGCTGCTTACGAGGTCTGAATGGGCTCC 1053

Qy 270 TTGGGAGGACGACAGG 286
    |||||
Db 1054 TTGATCGGCCAGACTGG 1070
```

RESULT 15

```
US-09-614-221A-399
; Sequence 399, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; PRIORITY FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 399
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-399
```

```
Query Match 4.0%; Score 38.8; DB 4; Length 1875;
Best Local Similarity 54.1%; Pred. No. 0.39;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 155 ACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGAC 214
Db 796 ACAAATATGCCGTCACGCCATCGATATTGCCAAAGATTTAGATATCAGCAAAATACGAT 855

Qy 215 GGCATCTCTGTGCGGGGAGATGTTTTCAGCGAGGTGTCGACGGTCTCATTTGGG 274
Db 856 ACCATTGCATGTGCTCGGGTGGTATTCCATACGAAGTAATTAATGGCTTTATAGA 915
```

Qy 275 AGGACGCAGAGAGCGCCGGGTCA 300
|||
Db 916 AGACCCGACAGAGTGGATGCGTTCAA 941
|||

Search completed: September 5, 2005, 15:12:11
Job time : 189.716 secs

This Page Blank (uspto)

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|-----|-----|-------------|
| 1 | 100 | 100 | 100 | 100 | 100 | 100 |

| | | | | | | |
|---|-----|-------|------|----|------------------|--------------------|
| 1 | 979 | 100.0 | 979 | 10 | US-09-969-896-1 | Sequence 1, Appli |
| 2 | 979 | 100.0 | 979 | 19 | US-10-631-958-1 | Sequence 1, Appli |
| 3 | 979 | 100.0 | 1840 | 9 | US-09-784-810A-5 | Sequence 5, Appli |
| 4 | 979 | 100.0 | 1840 | 22 | US-10-876-281-5 | Sequence 5, Appli |
| 5 | 903 | 92.2 | 1614 | 10 | US-09-969-896-9 | Sequence 9, Appli |
| 6 | 903 | 92.2 | 1614 | 19 | US-10-631-958-9 | Sequence 9, Appli |
| 7 | 903 | 92.2 | 1740 | 18 | US-10-262-511-39 | Sequence 39, Appli |

QY 781 GGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCAATAATGTCCTG 840
Db 781 GGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCAATAATGTCCTG 840
QY 841 TGTCTGTGCGCGGAGCGCCAGAGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTC 900
Db 841 TGTCTGTGCGCGGAGCGCCAGAGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTC 900
QY 901 TGACCTCATCTCTATCCGGAATGTCCAGGTTCAATTTCTGAGATTCTCATCAGCA 960
Db 901 TGACCTCATCTCTATCCGGAATGTCCAGGTTCAATTTCTGAGATTCTCATCAGCA 960
QY 961 CACCAACGAGGAGGACCAG 979
Db 961 CACCAACGAGGAGGACCAG 979
RESULT 3
US-09-784-810A-5
; Sequence 5, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-784-810A-5

Query Match 100.0%; Score 979; DB 9; Length 1840;
Best Local Similarity 100.0%; Pred. No. 1.4e-308;
Matches 979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGAAAGGACAAAGCAAGCGAT 60
Db 264 ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGAAAGGACAAAGCAAGCGAT 323
QY 61 ATATGAAAGAAAGTGGCACCCTGTTCACTTAGCCTCCATCACCCTGACATCATCGG 120
Db 324 ATATGAAAGAAAGTGGCACCCTGTTCACTTAGCCTCCATCACCCTGACATCATCGG 383
QY 121 TAACAAATTCATGTTAACTATGAGAAGTAATTAAGTAAGTCTGCTGCGCGGAGATGG 180
Db 384 TAACAAATTCATGTTAACTATGAGAAGTAATTAAGTAAGTCTGCTGCGCGGAGATGG 443
QY 181 GACTCTGATGAGTTAAATAGACAAATACGACGGCATGCTGCTGCGCGGAGATGG 240
Db 444 GACTCTGATGAGTTAAATAGACAAATACGACGGCATGCTGCTGCGCGGAGATGG 503
QY 241 TATGTTACGAGGTGCTGCACGGTCTGATTTGGGAGGACGACAGAGGACGCGCGGGTGA 300
Db 504 TATGTTACGAGGTGCTGCACGGTCTGATTTGGGAGGACGACAGAGGACGCGCGGGTGA 563
QY 301 CCAGAACACCCCGGGCTGCTGCTGCCAGTAGCTCCGGATTGGAATCATTTCCGC 360
Db 564 CCAGAACACCCCGGGCTGCTGCTGCCAGTAGCTCCGGATTGGAATCATTTCCGC 623
QY 361 AGGGTCAACGAGTCTGCTGTTACTCCACCGTGGGACACGACGACGAGAACTCGGC 420
Db 624 AGGGTCAACGAGTCTGCTGTTACTCCACCGTGGGACACGACGACGAGAACTCGGC 683

QY 421 GCTGATATCGTTGTTGGGACTGCTGCGCCATGATGTGTCTCAGTCCACCAACAG 480
Db 684 GCTGATATCGTTGTTGGGACTGCTGCGCCATGATGTGTCTCAGTCCACCAACAG 743
QY 481 CACACTCTCTCGCTACTCCGCTGCTGCTGGGCTTACGGCTTCTACGGGACATCATCA 540
Db 744 CACACTCTCTCGCTACTCCGCTGCTGCTGGGCTTACGGCTTCTACGGGACATCATCA 803
QY 541 GGACAGTGAGAGAAACGGTGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 600
Db 804 GGACAGTGAGAGAAACGGTGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 863
QY 601 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGACAAACACACGGT 660
Db 864 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGACAAACACACGGT 923
QY 661 GGGATCTCCAAGGATAGGAAGCCCTGCGCGGAGGATGCTTGTGTTTCAGGCAAGCAA 720
Db 924 GGGATCTCCAAGGATAGGAAGCCCTGCGCGGAGGATGCTTGTGTTTCAGGCAAGCAA 983
QY 721 GCAGCAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAGCTGGGAGGACGT 780
Db 984 GCAGCAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAGCTGGGAGGACGT 1043
QY 781 GGAGAGTGGCAAGTCTGTGGAAGTCTTGTGCGCATCAATGCCAATAATGTCCTG 840
Db 1044 GGAGAGTGGCAAGTCTGTGGAAGTCTTGTGCGCATCAATGCCAATAATGTCCTG 1103
QY 841 TGCTTGTGCGCGGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTC 900
Db 1104 TGCTTGTGCGCGGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTC 1163
QY 901 TGACCTCATCTCTATCCGGAATGTCCAGGTTCAATTTCTGAGATTCTCATCAGCA 960
Db 1164 TGACCTCATCTCTATCCGGAATGTCCAGGTTCAATTTCTGAGATTCTCATCAGCA 1223
QY 961 CACCAACGAGGAGGACCAG 979
Db 1224 CACCAACGAGGAGGACCAG 1242
RESULT 4
US-10-876-281-5
; Sequence 5, Application US/10876281
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-876-281-5

Query Match 100.0%; Score 979; DB 22; Length 1840;
Best Local Similarity 100.0%; Pred. No. 1.4e-308;
Matches 979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGAAAGGACAAAGCAAGCGAT 60
Db 264 ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGAAAGGACAAAGCAAGCGAT 323


```
Db 1134 GGAGAGTGGCAAGTCGTCTGTGGGAAGTTCTGGCCATCAATGCCAACAACATGCTCTG 1193
Qy 841 TGCCTTGTCCCGAGAGCCCGAGGGGCTCTCCCGGCTGCCCACTTTGGGAGACGGGTCTTC 900
Db 1194 TGCCTTGTCCCGAGAGCCCGAGGGGCTCTCCCGGCTGCCCACTTTGGGAGACGGGTCTTC 1253
Qy 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGCA 960
Db 1254 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGCA 1313
Qy 961 CACCAACAGCAGGACCAG 979
Db 1314 CACCAACAGCAGGACCAG 1332

RESULT 6
US-10-631-958-9
; Sequence 9, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-9
```

Query Match 92.2%; Score 903; DB 19; Length 1614;

Best Local Similarity 96.6%; Pred. No. 9.3e-284; Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;

```
Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGACAAGCAAGCGGAT 60
Db 387 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGACAAGCAAGCGGAT 446
Qy 61 ATATGAAAGAAAGTGGCAACCACTGTTCACTTACCTCCATCCACTGACATCATCGG 120
Db 447 ATATGAAAGAAAGTGGCAACCACTGTTCACTTACCTCCATCCACTGACATCATCG- 505
Qy 121 TAACAAATTCATGTTAACTATGTAGAGTAATTAAGTCAATGCTAATCAGGCCAAGGA 180
Db 506 -----TTACTGAACATGCTAATCAGGCCAAGGA 533
Qy 181 GACTCTGATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGCGCGGAGATGG 240
Db 534 GACTCTGATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGCGCGGAGATGG 593
Qy 241 TATGTTACGAGAGTGTGCAAGTCTGATTTGGGAGGACGACAGAGAGCGCGGGGTGCA 300
Db 594 TATGTTACGAGAGTGTGCAAGTCTGATTTGGGAGGACGACAGAGAGCGCGGGGTGCA 653
Qy 301 CCAGAACCAACCCCGGGCTGTGCTCGTCCCGAGTAGCTCCGGATTGGAATCATTCCTCGC 360
Db 654 CCAGAACCAACCCCGGGCTGTGCTCGTCCCGAGTAGCTCCGGATTGGAATCATTCCTCGC 713
Qy 361 AGGGTCAACGAGTCTCGTGTGTTACTTCCACCGTGGGACCCAGCAGCAGAAACCTCGGC 420
Db 714 AGGGTCAACGAGTCTCGTGTGTTACTTCCACCGTGGGACCCAGCAGCAGAAACCTCGGC 773
```

```
Qy 421 GCTCATATCTTGTGTGGGGACTCGTGGCCATGATGTGTCTCAGTCCACCACAACAG 480
Db 774 GCTCATATCTTGTGTGGGGACTCGTGGCCATGATGTGTCTCAGTCCACCACAACAG 833
Qy 481 CACACTCTTCCGCTACTCCGTTGCTCCGCTAGGGCTTCTACGGGGACATCATCAA 540
Db 834 CACACTCTTCCGCTACTCCGTTGCTCCGCTAGGGCTTCTACGGGGACATCATCAA 893
Qy 541 GGACAGTGAGAGAAACGGTGTGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 600
Db 894 GGACAGTGAGAGAAACGGTGTGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 953
Qy 601 CTTCTCTCCCAACCACTGCTATGAGGACAGTGTCTTCTCCTGCAACACACCGGT 660
Db 954 CTTCTCTCCCAACCACTGCTATGAGGACAGTGTCTTCTCCTGCAACACACCGGT 1013
Qy 661 GGGATCTCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTCAGGCAAGCAA 720
Db 1014 GGGATCTCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTCAGGCAAGCAA 1073
Qy 721 GCAGCAGCTGGAGGAGGAGCAGAAAGCACTGTATGTTTGGAGCTGCGGAGACGT 780
Db 1074 GCAGCAGCTGGAGGAGGAGCAGAAAGCACTGTATGTTTGGAGCTGCGGAGACGT 1133
Qy 781 GGAGAGTGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCACAACATGTCCTG 840
Db 1134 GGAGAGTGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCACAACATGTCCTG 1193
Qy 841 TGCTTGTCCCGAGAGCCCGAGGGCCCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTC 900
Db 1194 TGCTTGTCCCGAGAGCCCGAGGGCCCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTC 1253
Qy 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTTCAATTTTCTGAGATTCTCATCAGGCA 960
Db 1254 TGACCTCATCTCATCCGGAATGCTCCAGGTTTCAATTTTCTGAGATTCTCATCAGGCA 1313
Qy 961 CACCAACAGCAGGACCAG 979
Db 1314 CACCAACAGCAGGACCAG 1332
```

RESULT 7

US-10-262-511-39

; Sequence 39, Application US/10262511

; Publication No. US20040038223A1

; GENERAL INFORMATION:

```
; APPLICANT: Smithson, Glenna
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zernusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
```

APPLICANT: Agee, Michele L.
APPLICANT: Berghs, Constance
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-462C
CURRENT APPLICATION NUMBER: US/10/262,511
CURRENT FILING DATE: 2003-05-28
PRIOR FILING DATE: 2001-10-02
PRIOR FILING DATE: 2001-10-02
PRIOR FILING DATE: 2002-04-19
PRIOR FILING DATE: 2002-04-19
PRIOR FILING DATE: 2001-10-09
PRIOR FILING DATE: 2002-05-17
PRIOR FILING DATE: 2002-05-17
PRIOR FILING DATE: 2002-10-09
PRIOR FILING DATE: 2002-10-09
PRIOR FILING DATE: 2002-04-17
PRIOR FILING DATE: 2002-05-16
PRIOR FILING DATE: 2002-04-19
PRIOR FILING DATE: 2001-10-09
PRIOR FILING DATE: 2001-10-05
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 439
SOFTWARE: CuraseqList version 0.1
SEQ ID NO 39
LENGTH: 1740
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (76)..(1686)
US-10-262-511-39

Query Match 92.2%; Score 903; DB 18; Length 1740;
Best Local Similarity 96.6%; Pred. No. 9.6e-284; Indels 33; Gaps 1;
Matches 946; Conservative 0; Mismatches 0

Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGGACAAAGCGGAT 60
Db 462 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGGACAAAGCGGAT 521

Qy 61 ATATGAAGAAAGTGCGACCACTGTTCACTTACCTTCATCAACCACTGATCATCGG 120
Db 522 ATATGAAGAAAGTGCGACCACTGTTCACTTACCTTCATCAACCACTGATCATCG- 580

Qy 121 TAACAAATTCATGTTAACTATGTAGAAGTAATTACTGAACATGCTAATCAGGCCAAGGA 180
Db 581 -----TTACTGAACATGCTAATCAGGCCAAGGA 608

Qy 181 GACTCTGTATGAGATTAAACATAGACAAATACGCGCATCTGTGTGCGCGGAGATGG 240
Db 609 GACTCTGTATGAGATTAAACATAGACAAATACGCGCATCTGTGTGCGCGGAGATGG 668

Qy 241 TATGTTACGAGGTGCTGCAAGGTCTGATTGGAGGACGAGAGAGCGCGGGTCTGA 300
Db 669 TATGTTACGAGGTGCTGCAAGGTCTGATTGGAGGACGAGAGAGCGCGGGTCTGA 728

Qy 301 CCAGAACACCCCGGCTGTGCTGCCAGTAGCTCCGGATTGGAATCATTTCCCGC 360
Db 729 CCAGAACACCCCGGCTGTGCTGCCAGTAGCTCCGGATTGGAATCATTTCCCGC 788

Qy 361 AGGGTCAACGCACTGCTGTGTTACTCCACCGTGGGACCAAGCGACGAGAAACCTCGGC 420
Db 789 AGGGTCAACGCACTGCTGTGTTACTCCACCGTGGGACCAAGCGACGAGAAACCTCGGC 848

Qy 421 GCTGATATCGTTGTTGGGACTCGCTGGCCATGATGTGCTTCAGTCCACCAACAG 480
Db 849 GCTGATATCGTTGTTGGGACTCGCTGGCCATGATGTGCTTCAGTCCACCAACAG 908

Qy 481 CACACTCTTCGCTACTCCCGTCCCTGTGGGCTACGGCTTCTACGGGACATCATCAA 540
Db 909 CACACTCTTCGCTACTCCCGTCCCTGTGGGCTACGGCTTCTACGGGACATCATCAA 968

Qy 541 GGCAGTGAGAAAGAAACGGTGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGAC 600
Db 969 GGCAGTGAGAAAGAAACGGTGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGAC 1028

Qy 601 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTGCAACACACGGT 660
Db 1029 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTGCAACACACGGT 1088

Qy 661 GGGATCTTCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTGTGTTGAGGAAAGCAA 720
Db 1089 GGGATCTTCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTGTGTTGAGGAAAGCAA 1148

Qy 721 GCAGCAGCTGGAGGAGGAGAGAAAGCACTGTATGTTTGGAGCTCGGAGGACGT 780
Db 1149 GCAGCAGCTGGAGGAGGAGAGAAAGCACTGTATGTTTGGAGCTCGGAGGACGT 1208

Qy 781 GGCAGGTGCAAGTCTGTGCGGAAGTCTTGGCCATCAATGCCCAACACATGTCTG 840
Db 1209 GGCAGGTGCAAGTCTGTGCGGAAGTCTTGGCCATCAATGCCCAACACATGTCTG 1268

Qy 841 TGCTTGTGCGGAGCGCCAGGGCCCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTC 900
Db 1269 TGCTTGTGCGGAGCGCCAGGGCCCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTC 1328

Qy 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960
Db 1329 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1388

Qy 961 CACCAACACGACGACCCAG 979
Db 1389 CACCAACACGACGACCCAG 1407

RESULT 8
US-09-969-896-16
Sequence 16, Application US/09969896
Publication No. US20030125533A1
GENERAL INFORMATION:
APPLICANT: Kossida, Sophia
TITLE OF INVENTION: Regulation of human Sphingosine
FILE REFERENCE: 004974.00594
CURRENT APPLICATION NUMBER: US/09/969,896
CURRENT FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: US 60/238,005
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/314,113
PRIOR FILING DATE: 2001-08-23
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 4413
TYPE: DNA
ORGANISM: Homo sapiens
US-09-969-896-16

Query Match 92.2%; Score 903; DB 10; Length 4413;
Best Local Similarity 96.6%; Pred. No. 1.6e-283; Indels 33; Gaps 1;
Matches 946; Conservative 0; Mismatches 0

Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGGACAAAGCGGAT 60
Db 462 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGGACAAAGCGGAT 521

Qy 61 ATATGAAGAAAGTGCGACCACTGTTCACTTACCTTCATCACCACCTGATCATCGG 120
Db 522 ATATGAAGAAAGTGCGACCACTGTTCACTTACCTTCATCACCACCTGATCATCG- 580

```
OY 121 TAACAAATTCATGTTAACTATGTAGAGTAATTAATGACATGCTAATCAGGCCAAGGA 180
Db 581 -----TTACTGAACATGTAATCAGGCCAAGGA 608

OY 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGG 240
Db 609 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGG 668

OY 241 TATGTTTACGCGAGTGTGTCAGCGTCTGANTGGGAGGACGACAGAGAGCGCGCGGGTCTGA 300
Db 669 TATGTTTACGCGAGTGTGTCAGCGTCTGANTGGGAGGACGACAGAGAGCGCGCGGGTCTGA 728

OY 301 CCAGAACACACCCCGGGTGTGCTGCTCCAGTAGCCTCCGGATTGGAATCATCTCCGC 360
Db 729 CCAGAACACACCCCGGGTGTGCTGCTCCAGTAGCCTCCGGATTGGAATCATCTCCGC 788

OY 361 AGGTCACACGGAATCGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCGGC 420
Db 789 AGGTCACACGGAATCGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCGGC 848

OY 421 GCTGATATCGTGTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAG 480
Db 849 GCTGATATCGTGTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAG 908

OY 481 CACACTCTTCGCTACTCGTGTCTCCTGCTGGGCTACGGCTTCTACGGGACATCATCA 540
Db 909 CACACTCTTCGCTACTCGTGTCTCCTGCTGGGCTACGGCTTCTACGGGACATCATCA 968

OY 541 GGACAGTGAAGAAACGGTGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 600
Db 969 GGACAGTGAAGAAACGGTGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 1028

OY 601 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTGTCACCAACACAGGT 660
Db 1029 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTGTCACCAACACAGGT 1088

OY 661 GGGATCTCCAGGATAGAAACCGTCTGCGGCGAGGATGCTTTGTTCCAGGCAAGCAA 720
Db 1089 GGGATCTCCAGGATAGAAACCGTCTGCGGCGAGGATGCTTTGTTCCAGGCAAGCAA 1148

OY 721 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAGCTGCGGAGGAGCT 780
Db 1149 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAGCTGCGGAGGAGCT 1208

OY 781 GGAGGAGTGGCAAGTCTGTGGAAGTCTTGGCCATCAATGCCCAAAACATGTCTCTG 840
Db 1209 GGAGGAGTGGCAAGTCTGTGGAAGTCTTGGCCATCAATGCCCAAAACATGTCTCTG 1268

OY 841 TGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGAGGGTCTTC 900
Db 1269 TGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGAGGGTCTTC 1328

OY 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960
Db 1329 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1388

OY 961 CACCAACACGAGGACCCAG 979
Db 1389 CACCAACACGAGGACCCAG 1407
```

RESULT 9

```
US-10-631-958-16
; Sequence 16, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PENDING FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
```

```
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 4413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-16
```

Query Match 92.2%; Score 903; DB 19; Length 4413;

Best Local Similarity 96.6%; Pred. No. 1.6e-283;

Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;

```
OY 1 ACCAAGCATTTACTGTGTTATTTATCAACCCGTTTGGAGGAAAGGACAAAGCCAGCGAT 60
Db 462 ACCAAGCATTTACTGTGTTATTTATCAACCCGTTTGGAGGAAAGGACAAAGCCAGCGAT 521

OY 61 ATATGAAAGAAAGTGGCACCCACTGTTACCTTAGCCTTCCATCACCCTGACATCATCGG 120
Db 522 ATATGAAAGAAAGTGGCACCCACTGTTACCTTAGCCTTCCATCACCCTGACATCATCG- 580

OY 121 TAACAAAATTCATGTTAACTATGTAGAGTAATTAATGACATGCTAATCAGGCCAAGGA 180
Db 581 -----TTACTGAACATGCTAATCAGGCCAAGGA 608

OY 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGG 240
Db 609 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGG 668

OY 241 TATGTTTACGCGAGTGTGTCAGCGTCTGANTGGGAGGACGACAGAGGAGCGCGGGTCTGA 300
Db 669 TATGTTTACGCGAGTGTGTCAGCGTCTGANTGGGAGGACGACAGAGGAGCGCGGGTCTGA 728

OY 301 CCAGAACACACCCCGGGTGTGTCGCTCCAGTAGCCTTCCGGATTTGGAATCATTTCCGC 360
Db 729 CCAGAACACACCCCGGGTGTGTCGCTCCAGTAGCCTTCCGGATTTGGAATCATTTCCGC 788

OY 361 AGGTCACACGGAATCGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCGGC 420
Db 789 AGGTCACACGGAATCGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCGGC 848

OY 421 GCTGATATCGTGTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAG 480
Db 849 GCTGATATCGTGTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAG 908

OY 481 CACACTCTTCGCTACTCGTGTCTCCTGCTGGGCTACGGCTTCTACGGGACATCATCA 540
Db 909 CACACTCTTCGCTACTCGTGTCTCCTGCTGGGCTACGGCTTCTACGGGACATCATCA 968

OY 541 GGACAGTGAAGAAACGGTGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 600
Db 969 GGACAGTGAAGAAACGGTGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 1028

OY 601 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTGTCACCAACACAGGT 660
Db 1029 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTGTCACCAACACAGGT 1088

OY 661 GGGATCTCCAGGATAGAAACCGTCTGCGGCGAGGATGCTTTGTTCCAGGCAAGCAA 720
Db 1089 GGGATCTCCAGGATAGAAACCGTCTGCGGCGAGGATGCTTTGTTCCAGGCAAGCAA 1148

OY 721 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAGCTGCGGAGGAGCT 780
Db 1149 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAGCTGCGGAGGAGCT 1208

OY 781 GGAGGAGTGGCAAGTCTGTGGAAGTCTTGGCCATCAATGCCCAAAACATGTCTCTG 840
Db 1209 GGAGGAGTGGCAAGTCTGTGGAAGTCTTGGCCATCAATGCCCAAAACATGTCTCTG 1268

OY 841 TGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGAGGGTCTTC 900
Db 1269 TGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGAGGGTCTTC 1328

OY 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960
Db 1329 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1388

OY 961 CACCAACACGAGGACCCAG 979
Db 1389 CACCAACACGAGGACCCAG 1407
```

Qy 841 TGCTTGTCCGCGAGCCCGGGGCTCTCCCGCTGCCCCACCTGGGAGACGGGTCTTC 900
Db 1269 TGCTTGTCCGCGAGCCCGGGGCTCTCCCGCTGCCCCACCTGGGAGACGGGTCTTC 1328
Qy 901 TGACCTCATCTCATCCGGAATGCTCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 960
Db 1329 TGACCTCATCTCATCCGGAATGCTCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 1388
Qy 961 CACCAACAGCAGGACCAG 979
Db 1389 CACCAACAGCAGGACCAG 1407
RESULT 10
US-10-618-941-55
; Sequence 55, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: MANNING, DAVID
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 55
; LENGTH: 4429
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-618-941-55

Query Match 92.2%; Score 903; DB 19; Length 4429;
Best Local Similarity 96.6%; Pred. No. 1.6e-283;
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;
Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGACAGCAAGCGGAT 60
Db 478 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGACAGCAAGCGGAT 537
Qy 61 ATATGAAAGAAAGTGGCACCACCTGTTCACTTAGCTTCCATCACCCTGACATCATCGG 120
Db 538 ATATGAAAGAAAGTGGCACCACCTGTTCACTTAGCTTCCATCACCCTGACATCATCG- 596
Qy 121 TAACAAATTTCTATGTTAACTATGTAGAAAGTAACTACTGAAACATGCTAATCAGGCCAAGGA 180
Db 597 -----TTACTGAACATGCTAATCAGGCCAAGGA 624
Qy 181 GACTCTGTATGATTAACATAGACAAATACGCGGCATCTCTGTGTCGGCGGAGATGG 240
Db 625 GACTCTGTATGATTAACATAGACAAATACGCGGCATCTCTGTGTCGGCGGAGATGG 684
Qy 241 TATGTTACGAGGTGCTGCAGGTCTGATTGGGAGGACGACAGAGCGCGGGGTGCA 300
Db 685 TATGTTACGAGGTGCTGCAGGTCTGATTGGGAGGACGACAGAGCGCGGGGTGCA 744
Qy 301 CCAGAACACCCCGGGGTGTGCTGGTCCCAGTAGCTCCCGGATTTGGAATCATTTCCGCG 360
Db 745 CCAGAACACCCCGGGGTGTGCTGGTCCCAGTAGCTCCCGGATTTGGAATCATTTCCGCG 804
Qy 361 AGGCTCAAGCACTCGGTGTTTACTCCACCGTGGGACCGACGCGAGAAACCTCGGC 420
Db 805 AGGCTCAAGCACTCGGTGTTTACTCCACCGTGGGACCGACGCGAGAAACCTCGGC 864
Qy 421 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCCTCAGTCCACCAACAG 480
Db 865 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCCTCAGTCCACCAACAG 924
Qy 481 CACACTCTCGTACTCTCGGTGCTCGGTGCTGAGGCTTACCGGGACATCATCA 540

Db 925 CACACTCTTCGCTACTCCGTCCTGCTGGGCTACTACGGCTTCTTACGGGACATCATCAA 984
Qy 541 GGCAGTGAAGAAACGGTGGTCTTGGCTCTTGGCAGATACGACTTTTTCAGGTTTAAAGAC 600
Db 985 GGCAGTGAAGAAACGGTGGTCTTGGCTCTTGGCAGATACGACTTTTTCAGGTTTAAAGAC 1044
Qy 601 CTTTCTCTCCACACCTGCTATGAAAGGACAGTGTCTCTCTCCCTGCACAACACACCGGT 660
Db 1045 CTTTCTCTCCACACCTGCTATGAAAGGACAGTGTCTCTCTCCCTGCACAACACACCGGT 1104
Qy 661 GGGATCTCAAGGATAGGAGCCCTCCCGGGCAGGATGCTTTGTTTGGAGGAAAGCAA 720
Db 1105 GGGATCTCAAGGATAGGAGCCCTCCCGGGCAGGATGCTTTGTTTGGAGGAAAGCAA 1164
Qy 721 GCAGCAGCTGGAGGAGCAGAGAAAGCACTGTATGTTTGGAAAGCTTGGAAAGCTG 780
Db 1165 GCAGCAGCTGGAGGAGCAGAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGGACGT 1224
Qy 781 GGAGGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTG 840
Db 1225 GGAGGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTG 1284
Qy 841 TGCTTGTCCGCGAGAGCCCGAGGGGCTCTCCCGGCTGCGCCACTTGGGAGACGGGTCTTC 900
Db 1285 TGCTTGTCCGCGAGAGCCCGAGGGGCTCTCCCGGCTGCGCCACTTGGGAGACGGGTCTTC 1344
Qy 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960
Db 1345 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1404
Qy 961 CACCAACAGCAGGACCAG 979
Db 1405 CACCAACAGCAGGACCAG 1423

RESULT 11
US-10-988-148
; Sequence 148, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radojic T.
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 148
; LENGTH: 4432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129) .. (2817)
US-10-988-148

Query Match 92.1%; Score 901.4; DB 17; Length 4432;
Best Local Similarity 96.5%; Pred. No. 5.3e-283;
Matches 945; Conservative 0; Mismatches 1; Indels 33; Gaps 1;
Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGGACAAAGCAAGCGGAT 60
Db 1590 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGGACAAAGCAAGCGGAT 1649
Qy 61 ATATGAAAGAAAGTGGACCCACTGTTTCACTTAGCCTTCCATCACCCTGACATCATCGG 120


```
Qy 781 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAAATGTCCTG 840
Db 1257 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAAATGTCCTG 1316

Qy 841 TGCTTTGTCGCGGAGCCCCAGGGGCTTCTCCCGGCTGCCCCACATCTGGGAGACGGGTCTTC 900
Db 1317 TGCTTTGTCGCGGAGCCCCAGGGGCTTCTCCCGGCTGCCCCACATCTGGGAGACGGGTCTTC 1376

Qy 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCA 960
Db 1377 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCA 1436

Qy 961 CACCAACAGCAGGACCAG 979
Db 1437 CACCAACAGCAGGACCAG 1455

RESULT 13
US-09-969-896-4
; Sequence 4, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-4

Query Match 47.4%; Score 464.4; DB 10; Length 474;
Best Local Similarity 99.8%; Pred. No. 1.2e-140;
Matches 465; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 239 GGTATGTTACGAGGTCTGCACGGTCTGATTGGAGGACGACAGAGCGCGGGGTC 298
Db 8 GGTATGTTACGAGGTCTGCACGGTCTGATTGGAGGACGACAGAGCGCGGGGTC 67

Qy 299 GACCAGAACCAACCCCGGGCTGTGCTGCTCCCAAGTAGCCTCCGGATTGGAATCATTTCCC 358
Db 68 GACCAGAACCAACCCCGGGCTGTGCTGCTCCCAAGTAGCCTCCGGATTGGAATCATTTCCC 127

Qy 359 GCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCG 418
Db 128 GCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCG 187

Qy 419 GCGTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCAGTCCACCAACAAC 478
Db 188 GCGTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCAGTCCACCAACAAC 247

Qy 479 AGCACACTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATC 538
Db 248 AGCACACTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATC 307

Qy 539 AAGGACAGTGAGAGAAACCGGTGGTCTTGCGGCTTTGCGCAGATACGACTTTTCAGGTTTAAAG 598
Db 128 GCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCG 187

Qy 419 GCGTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCAGTCCACCAACAAC 478
Db 188 GCGTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCAGTCCACCAACAAC 247

Qy 479 AGCACACTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATC 538
Db 248 AGCACACTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATC 307

Qy 539 AAGGACAGTGAGAGAAACCGGTGGTCTTGCGGCTTTGCGCAGATACGACTTTTCAGGTTTAAAG 598
Db 308 AAGGACAGTGAGAGAAACCGGTGGTCTTGCGCAGATACGACTTTTCAGGTTTAAAG 367

Qy 599 ACCTTCTCTCCACCACTGCTATGAGGGACAGTGTCTTCTCCCTGCGACACACACG 658
Db 368 ACCTTCTCTCCACCACTGCTATGAGGGACAGTGTCTTCTCCCTGCGACACACACG 427

Qy 659 GTGGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTG 704
```

```
Db 428 GTGGGATCTCCAAGGATAGGAAGCCCTGCGGGCAAGATGCTTTG 473

RESULT 14
US-10-631-958-4
; Sequence 4, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-4

Query Match 47.4%; Score 464.4; DB 19; Length 474;
Best Local Similarity 99.8%; Pred. No. 1.2e-140;
Matches 465; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 239 GGTATGTTACGAGGTCTGCACGGTCTGATTGGAGGACGACAGAGCGCGGGGTC 298
Db 8 GGTATGTTACGAGGTCTGCACGGTCTGATTGGAGGACGACAGAGCGCGGGGTC 67

Qy 299 GACCAGAACCAACCCCGGGCTGTGCTGCTCCCAAGTAGCCTCCGGATTGGAATCATTTCCC 358
Db 68 GACCAGAACCAACCCCGGGCTGTGCTGCTCCCAAGTAGCCTCCGGATTGGAATCATTTCCC 127

Qy 359 GCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCG 418
Db 128 GCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCG 187

Qy 419 GCGTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCAGTCCACCAACAAC 478
Db 188 GCGTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCAGTCCACCAACAAC 247

Qy 479 AGCACACTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATC 538
Db 248 AGCACACTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATC 307

Qy 539 AAGGACAGTGAGAGAAACCGGTGGTCTTGCGGCTTTGCGCAGATACGACTTTTCAGGTTTAAAG 598
Db 308 AAGGACAGTGAGAGAAACCGGTGGTCTTGCGCAGATACGACTTTTCAGGTTTAAAG 367

Qy 599 ACCTTCTCTCCACCACTGCTATGAGGGACAGTGTCTTCTCCCTGCGACACACACG 658
Db 368 ACCTTCTCTCCACCACTGCTATGAGGGACAGTGTCTTCTCCCTGCGACACACACG 427

Qy 659 GTGGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTG 704
Db 428 GTGGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTG 473

RESULT 15
US-09-969-896-5
; Sequence 5, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
```



```
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(329)
; OTHER INFORMATION: n = A,T,C or G
US-09-969-896-5
```

```
Query Match          30.9%; Score 302.2; DB 10; Length 329;
Best Local Similarity 97.7%; Pred. No. 1e-87;
Matches 304; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 153 TTACTGAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACG 212
Db 19 TTACTGAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACG 78

QY 213 ACGGCATCGTCTGTGTCGCGGAGATGGTATGTTTCAGCGAGGTGCTGCACGGTCTGATTG 272
Db 79 ACGGCATCGTCTGTGTCGCGGAGATGGTATGTTTCAGCGAGGTGCTGCACGGTCTGATTG 138

QY 273 GGAGACGACAGAGGAGCGCGGGGTGACACAGAACACCCCGGGCTGTGCTGGTCCCA 332
Db 139 GGAGACGACAGAGGAGCGCGGGGTGACACAGAACACCCCGGGCTGTGCTGGTCCCA 198

QY 333 GTAGCCTCGGATTGGAATCATTCGCGAGGTCAGCGACTGCGTGTGTTACTCCACG 392
Db 199 GTAGCCTCGGATTGGAATCATTCGCGAGGTCAGCGACTGCGTGTGTTACTCCACG 258

QY 393 TGGGCACGACGACGACGACAAACCTCGGCGTGCATATCGTTGTTGGGACTCGCTGGCCA 452
Db 259 TGGGCACGACGACGACGACAAACCTCGGCGTGCATATCGTTGTTGGGACTCGCTGGCCA 318

QY 453 TGGATGTGTCC 463
Db 319 TGGATGTGTCC 329
```

Search completed: September 5, 2005, 20:00:08
Job time : 726.691 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 18:18:34 ; Search time 3579.6 Seconds
(without alignments)
10410.363 Million cell updates/sec

Title: US-10-631-958-1

Perfect score: 979
Sequence: 1 accaaagattactgtat.....acaccaaccagcaggaccag 979

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_hc:*
 - 4: gb_est3:*
 - 5: gb_est4:*
 - 6: gb_est5:*
 - 7: gb_est6:*
 - 8: gb_gss1:*
 - 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 801.4 | 81.9 | 1063 | BM479389 | BM479389 AGENCOURT |
| 2 | 726.4 | 74.2 | 1059 | BQ054406 | BQ054406 AGENCOURT |
| 3 | 696.6 | 71.2 | 797 | CK000755 | CK000755 AGENCOURT |
| 4 | 670 | 68.4 | 1047 | BQ057191 | BQ057191 AGENCOURT |
| 5 | 635.8 | 64.9 | 1824 | AK042077 | AK042077 Mus muscu |
| 6 | 635.8 | 64.9 | 4248 | AK052269 | AK052269 Mus muscu |
| 7 | 593.4 | 60.6 | 820 | CD655311 | CD655311 AGENCOURT |
| 8 | 582.4 | 59.5 | 1078 | BQ063738 | BQ063738 AGENCOURT |
| 9 | 579.4 | 59.2 | 581 | BP224560 | BP224560 BP224560 |
| 10 | 537.4 | 54.9 | 758 | CB246749 | CB246749 UI-M-F10- |
| 11 | 536.2 | 54.8 | 584 | BP310011 | BP310011 BP310011 |
| 12 | 535.8 | 54.7 | 541 | CN296311 | CN296311 170005321 |
| 13 | 527.4 | 53.9 | 584 | BP309990 | BP309990 BP309990 |
| 14 | 520.8 | 53.2 | 573 | CF138275 | CF138275 UI-HF-BNO |
| 15 | 518.2 | 52.9 | 550 | CF138634 | CF138634 UI-HF-BNO |
| 16 | 505 | 51.6 | 713 | CK000525 | CK000525 AGENCOURT |
| 17 | 503 | 51.4 | 661 | CN296312 | CN296312 170005321 |
| 18 | 477.2 | 48.7 | 488 | BW430459 | BW430459 UI-HF-BNO |
| 19 | 464.4 | 47.4 | 474 | AW503999 | AW503999 UI-HF-BNO |
| 20 | 445 | 45.5 | 653 | BX952302 | BX952302 DKFZp781L |
| 21 | 415 | 42.4 | 732 | CF135528 | CF135528 UI-HF-BNO |
| 22 | 402.8 | 41.1 | 576 | CB608292 | CB608292 AMGNNUC:N |
| 23 | 367.8 | 37.6 | 545 | CB611947 | CB611947 AMGNNUC:N |
| 24 | 367.4 | 37.5 | 376 | BU430866 | BU430866 UI-HF-BNO |

| | | | | | | |
|----|-------|------|-----|---|----------|--------------------|
| 25 | 348.8 | 35.6 | 359 | 5 | BU428926 | BU428926 UI-HF-BNO |
| 26 | 340 | 34.7 | 553 | 6 | CA578972 | CA578972 K0731A08- |
| 27 | 318.8 | 32.6 | 725 | 5 | BU214294 | BU214294 603756937 |
| 28 | 306.2 | 31.3 | 674 | 1 | AJ739034 | AJ739034 AJ739034 |
| 29 | 302.2 | 30.9 | 329 | 1 | AA355581 | AA355581 EST64087 |
| 30 | 294 | 30.0 | 779 | 5 | BU265377 | BU265377 60374309 |
| 31 | 288.6 | 29.5 | 467 | 4 | BM484519 | BM484519 538413 MA |
| 32 | 278.8 | 28.5 | 763 | 7 | CK364327 | CK364327 AGENCOURT |
| 33 | 278.6 | 28.5 | 385 | 5 | BU428707 | BU428707 UI-HF-BNO |
| 34 | 274.8 | 28.1 | 902 | 5 | BU373554 | BU373554 603811294 |
| 35 | 268.2 | 27.4 | 607 | 6 | CB581157 | CB581157 AMGNNUC:N |
| 36 | 267.2 | 27.3 | 452 | 6 | CA579952 | CA579952 K0744G05- |
| 37 | 263.2 | 26.9 | 735 | 7 | CF535441 | CF535441 UI-M-GH0- |
| 38 | 260 | 26.6 | 310 | 6 | CD631568 | CD631568 56059637H |
| 39 | 258.4 | 26.4 | 310 | 6 | CD631566 | CD631566 56059545H |
| 40 | 253.6 | 25.9 | 759 | 5 | BU290807 | BU290807 604162346 |
| 41 | 242 | 24.7 | 291 | 6 | CD631569 | CD631569 56059637J |
| 42 | 239.8 | 24.5 | 292 | 6 | CD631567 | CD631567 56059545J |
| 43 | 234.4 | 23.9 | 759 | 5 | BU264162 | BU264162 603814452 |
| 44 | 222 | 22.7 | 567 | 1 | AJ739023 | AJ739023 AJ739023 |
| 45 | 215.6 | 22.0 | 638 | 1 | AJ739022 | AJ739022 AJ739022 |

ALIGNMENTS

RESULT 1
BM479389
LOCUS BM479389 1063 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT 6418742 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5502281
5', mRNA sequence.
ACCESSION BM479389
VERSION BM479389.1 GI:18528431
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1063)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12140 row: c column: 18
High quality sequence stop: 665.
FEATURES
Location/Qualifiers
1..1063
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5502281"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Library constructed by Life Technologies."

ORIGIN
Query Match 81.9%; Score 801.4; DB 4; Length 1063;
Best Local Similarity 98.7%; Pred. No. 5.2e-219;
Matches 805; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 153 TTACTGAACATGCTAATCAGCCAGGAGACTCTGTATGAGATTACATACAAATACG 212

| | | | | |
|-------------------------------------------------------------------|-----|--------------------------------------------------------------|-----|--|
| Db | 44 | TTACTGAACATGCTAAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACG | 103 | |
| Qy | 213 | ACGGATCGTCTGTGTCGGCGAGATGTAATGTTTCAGGAGGCTGTCACGGTCTGATTG | 272 | |
| Db | 104 | ACGGATCGTCTGTGTCGGCGAGATGTAATGTTTCAGGAGGCTGTCACGGTCTGATTG | 163 | |
| Qy | 273 | GGAGAGCAGAGGAGCGCGGGTCGACACAGAACCCCGCGGCTGTGTCGGTCCCA | 332 | |
| Db | 164 | GGAGAGCAGAGGAGCGCGGGTCGACACAGAACCCCGCGGCTGTGTCGGTCCCA | 223 | |
| Qy | 333 | GTAGCTCCGATTGGAAATCAATCCCGCAGGGTCAACGGAATGCGTGTGTTACTCCACCG | 392 | |
| Db | 224 | GTAGCTCCGATTGGAAATCAATCCCGCAGGGTCAACGGAATGCGTGTGTTACTCCACCG | 283 | |
| Qy | 393 | TGGGACAGGAGCAGAGAAACCTCGGGCTGCAATATGTTGTTGGGACTGCTGTGGCCA | 452 | |
| Db | 284 | TGGGACAGGAGCAGAGAAACCTCGGGCTGCAATATGTTGTTGGGACTGCTGTGGCCA | 343 | |
| Qy | 453 | TGGATGTCTCTCAGTCCACACACAGCACTCTTCCGCTACTCCGCTGCTGTGCTGG | 512 | |
| Db | 344 | TGGATGTCTCTCAGTCCACACACAGCACTCTTCCGCTACTCCGCTGCTGTGCTGG | 403 | |
| Qy | 513 | GCTAGGGTCTTACGGGACATCATCAAGGACAGTGAGAGAAACGGTGTGGGTCTTG | 572 | |
| Db | 404 | GCTAGGGTCTTACGGGACATCATCAAGGACAGTGAGAGAAACGGTGTGGGTCTTG | 463 | |
| Qy | 573 | CCAGATAGCACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAGGGACAG | 632 | |
| Db | 464 | CCAGATAGCACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAGGGACAG | 523 | |
| Qy | 633 | TGTCCTCTCCCTGCACACACAGGCTGGGATCTTCAAGGATAGGAAGCCCTGCCGG | 692 | |
| Db | 524 | TGTCCTCTCCCTGCACACACAGGCTGGGATCTTCAAGGATAGGAAGCCCTGCCGG | 583 | |
| Qy | 693 | CAGGATGCTTTGTTGAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCAC | 752 | |
| Db | 584 | CAGGATGCTTTGTTGAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCAC | 643 | |
| Qy | 753 | TGATGTTTGGAGCTCGGAGACGTGGAGGAGTGCAAGTCTGTGTTGGGAGTTTC | 812 | |
| Db | 644 | TGATGTTTGGAGCTCGGAGACGTGGAGGAGTGCAAGTCTGTGTTGGGAGTTTC | 703 | |
| Qy | 813 | TGGCCATCAATGCCAACAAATGCTGTGCTGTGTCGGCGAGCCCAAGGCGCTCTCC | 872 | |
| Db | 704 | TGGCNATCATGCCAACAAATGCTGTGCTGTGTCGGCGAGCCCAAGGCGCTCTCNC | 763 | |
| Qy | 873 | CGGTGCCCCACTTGGGAGACGGTCTTCTGACCTCATCTTCATCCGAAATGCTCCAGGT | 932 | |
| Db | 764 | CGGTGCCCCACTTGGGAGACGGTCTTCTGACCTCATCTTCATCCGAAATGCTCCAGGN | 823 | |
| Qy | 933 | TCAATTTTCTGAGATTCTCATCAGGACACCAACC | 968 | |
| Db | 824 | TCAATTTTCTGAGATTCTCATCAGGACACCAACC | 859 | |
| RESULT 2 | | | | |
| BQ054406 | | | | |
| LOCUS | | | | |
| DEFINITION | | | | |
| AGENCOURT_6771237 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:580368 | | | | |
| 5', mRNA sequence. | | | | |
| BQ054406 | | | | |
| ACCESSION | | | | |
| BQ054406.1 GI:19813746 | | | | |
| KEYWORDS | | | | |
| EST. | | | | |
| SOURCE | | | | |
| Homo sapiens (human) | | | | |
| ORGANISM | | | | |
| Homo sapiens | | | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| 1 (bases 1 to 1059) | | | | |
| NIH-MGC http://mgc.nci.nih.gov/. | | | | |
| AUTHORS | | | | |
| National Institutes of Health, Mammalian Gene Collection (MGC) | | | | |
| JOURNAL | | | | |
| Unpublished (1999) | | | | |
| COMMENT | | | | |
| Contact: Robert Strausberg, Ph.D. | | | | |

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2040 row: a column: 13
High quality sequence stop: 601.
Location/Qualifiers
1. .1059
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:580368"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

| | | | | | |
|-----------------------|-----|-----------------------------------------------------------------|-------------------|-------|------------------------|
| Query Match | | 74.2%; | Score 726.4; | DB 5; | Length 1059; |
| Best Local Similarity | | 95.1%; | Pred. No. 2e-197; | | |
| Matches | | 784; | Conservative | 0; | Mismatches 34; Gaps 2; |
| Qy | 9 | ATTTACTGTTATTATCAACCCGTTTGGAGAAAGACAAGCGACGGATATATGAAA | 68 | | |
| Db | 13 | ATTTACTGTTATTATCAACCCGTTTGGAGAAAGACAAGCGACGGATATATGAAA | 72 | | |
| Qy | 69 | GAAAAGTGGACACCTGTTTCCACCTTAGCCCTTCCACTGACATCATCGTAACAAAT | 128 | | |
| Db | 73 | GAAAAGTGGACACCTGTTTCCACCTTAGCCCTTCCACTGACATCATCGT----- | 123 | | |
| Qy | 129 | TCTATGTTAACTATGTAGAAGTAATTAAGTAACTGTAATCAGGCGGAGAGACTCTGT | 188 | | |
| Db | 124 | -----TTACTGAACATGCTAATCAGGCCAAGGAGACTCTGT | 159 | | |
| Qy | 189 | ATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGTTATGTTCA | 248 | | |
| Db | 160 | ATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGTTATGTTCA | 219 | | |
| Qy | 249 | GCGAGGTGTCGACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGGTGACCCAGAAC | 308 | | |
| Db | 220 | GCGAGGTGTCGACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGGTGACCCAGAAC | 279 | | |
| Qy | 309 | ACCCCGGCGTGTGCTGCTGCTCCCGAGTAGCCCTCCGGATTGGAATCATTTCCCGAGGGTCAA | 368 | | |
| Db | 280 | ACCCCGGCGTGTGCTGCTGCTCCCGAGTAGCCCTCCGGATTGGAATCATTTCCCGAGGGTCAA | 339 | | |
| Qy | 369 | CGGACTCGTGTGTTACTTCCACCGTGGGACACGACGAGAAACCTCGGCGCTGCATA | 428 | | |
| Db | 340 | CGGACTCGTGTGTTACTTCCACCGTGGGACACGACGAGAAACCTCGGCGCTGCATA | 399 | | |
| Qy | 429 | TCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAGCACACTCC | 488 | | |
| Db | 400 | TCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAGCACACTCC | 459 | | |
| Qy | 489 | TTGCTACTCCGTGCTCCCTGCTGGGCTACCGGCTTCTACGGGACATCATCAAGGACAGTG | 548 | | |
| Db | 460 | TTGCTACTCCGTGCTCCCTGCTGGGCTTCTACGGGACATCATCAAGGACAGTG | 519 | | |
| Qy | 549 | AGAAAGAACGGTGGTGGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCT | 608 | | |
| Db | 520 | AGAAAGAACGGTGGTGGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCT | 579 | | |

QY 609 CCCACACTGCTATGAGGACAGTGTCTCTCTCCCTGCACAAACACACGCTGGATCTC 668
|||||
Db 580 CCCACACTGCTATGAACGACAGTGTCTCTCTCCCTGCACAAACACACGCTGGATCTC 639
QY 669 CAAGGATAGGAAGCCCTCCGGGAGGAGTGTCTCTCTCCCTGCAGGCAAGCAGCAGC 728
Db 640 CAAGGATAGGAAGCCCTCCGGGAGGAGTGTCTCTCTCCCTGCAGGCAAGCAGCAGC 699
QY 729 TGGAGGAGGAGCAGAAAGACACTGTATGGTTTGGAAAGCTCGGAGGACCTGGAGGAGT 788
Db 700 TGGAGGAGGAGCAGAAAGACACTGTATGGTTTGGAAAGCTCGGAGGAGCCTGGAGGAGT 759
QY 789 GCGAAGTCTCTGTGGG-AAGTTTCTGGCCATCATGCCCCAAA 831
Db 760 GCGAAGTCTCTGTGGGAAGTTCTGTGCCATTAATGCCACAAA 803

RESULT 3
CK000755
LOCUS
DEFINITION AGENCOURT_16363467 NIH_MGC_221 Homo sapiens cDNA clone
IMAGE:30707875 5', mRNA sequence.
ACCESSION CK000755
VERSION CK000755.1 GI:38526789
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 797)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: gcgabs-r@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM1073 row: c column: 20
High quality sequence stop: 656.
Location/Qualifiers
1..797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30707875"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_221"
/note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;
Site 2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. Average insert size
4-5Kb. Adaptors 5'(AATTCGACGAGG)3' and 5'd
(CTGTGTCGG)3'. 3' Linker sequence - CGGCGCGTGGAGCC T18.
Sequencing primers 3'end: T3 promoter primer 5'd
(ATTAACTCTCAATAAGGA)3'. 5' End: T7 promoter primer 5'd
(TAATACGACTCATATAGG)3'. Library was constructed in the
laboratory of M. Bento Soares. Note: this is a NIH_MGC
Library"

FEATURES
source

1047 bp mRNA linear EST 29-MAR-2002
BQ057191
AGENCOURT_6769595 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5812382
5', mRNA sequence.
ACCESSION BQ057191
VERSION BQ057191.1 GI:19816531
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1047)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

ORIGIN

Query Match 71.2%; Score 696.6; DB 7; Length 797;
Best Local Similarity 97.4%; Pred. No. 7e-189;
Matches 705; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 153 TTACTGAAATGCTTAATCAGGCCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACG 212
Db 61 TTACTGAAATGCTTAATCAGGCCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACG 120
QY 213 ACGCATCGTCTGTGTGGCGGAGATGATATGTTTTCAGCAGAGTGTCTGCACGGTCTGATTG 272
Db 121 ACGCATCGTCTGTGTGGCGGAGATGATATGTTTTCAGCAGAGTGTCTGCACGGTCTGATTG 180
QY 273 GGAGGACGACAGAGCGCGGGGTTCACACAGAACACCCCGGGCTGTGTGTCCTCCCA 332
Db 181 GGAGGACGACAGAGCGCGGGGTTCACACAGAACACCCCGGGCTGTGTGTCCTCCCA 240
QY 333 GTAGCTCCGATTTGGAAATCATTTCCGAGGGTCAACGAGCTGCGTGTGTTTACTTCCACCG 392
Db 241 GTAGCTCCGATTTGGAAATCATTTCCGAGGGTCAACGAGCTGCGTGTGTTTACTTCCACCG 300
QY 393 TGGGACACGAGCGACGAGAAACCTTCGGCGTGCATATGTTTGTGGGACTCGCTGGCCA 452
Db 301 TGGGACACGAGCGACGAGAAACCTTCGGCGTGCATATGTTTGTGGGACTCGCTGGCCA 360
QY 453 TGGATGTCTCTCAGTCCACACACAGCAGTCTCTCGCTACTCCGTGTCTCTGCTGG 512
Db 361 TGGATGTCTCTCAGTCCACACACAGCAGTCTCTCGCTACTCCGTGTCTCTGCTGG 420
QY 513 GCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACCGTGTGGTCTTTG 572
Db 421 GCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACCGTGTGGTCTTTG 480
QY 573 CCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGTATGAAGGACAG 632
Db 481 CCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGTATGAAGGACAG 540
QY 633 TGTCTCTCTCTCCGACACACACAGTGGGATCTCCAGGGATAGGAGCCCTGCCGG 692
Db 541 TGTCTCTCTCTCCGACACACACAGTGGGATCTCCAGGGATAGGAGCCCTGCCGG 600
QY 693 CAGGATGCTTTTGGAGCAAGCAAGCAGCAGTGGAGGAGGAGCAGAGAAAGCAC 752
Db 601 CAGGATGCTTTTGGAGCAAGCAAGCAGCAGTGGAGGAGGAGCAGAGAAAGCAC 660
QY 753 TGTATGTTTGGAAAGCTCGGAGGACGTTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTC 812
Db 661 TGTATGTTTGGAAAGCTCGGAGGACGTTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTC 720
QY 813 TGGCCATCAATGCCACAAACATGTCTGTCTCGCGGAGGCCCGGAGGCTCTCC 872
Db 721 TGGGCCNATCATGCAAAACATGTCTGTCTCGCGGAGGCCCGGAGGCTCTCCCGG 780
QY 873 CGCG 876
Db 781 CTGC 784

RESULT 4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2062 row: 1 column: 15
High quality sequence stop: 535.
Location/Qualifiers
1. 1047
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5812382"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 68.4%; Score 670; DB 5; Length 1047;
Best Local Similarity 91.3%; Pred. No. 3.5e-181;
Matches 825; Conservative 0; Mismatches 35; Indels 44; Gaps 9;
9 ATTTACTGGTATTTATCAACCGTTTGGAGAAAGGACGAAGCGGATATATGAAA 68
11 ATTTACTGGTATTTATCAACCGTTTGGAGAAAGGACGAAGCGGATATATGAAA 70
69 GAAAGTGGCACCAGTCTTACCTTAGCTTCATCACCAGTGCATCATCGGTAAACAAT 128
71 GAAAGTGGCACCAGTCTTACCTTAGCTTCATCACCAGTGCATCATCGGTAAACAAT 121
129 TCTATGTTAACTATGTAGAAAGTAACTACTGAACATGCTAATCAGCGCAAGGAGACTCTGT 188
122 -----TTACTGAACATGCTAATCAGCGCAAGGAGACTCTGT 157
189 ATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGTATGTCA 248
158 ATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGTATGTCA 217
249 GCGAGGTGCTGCACGGTCTGATTTGGGAGGACGACGAGGAGCGCGGGTGCACCAAGACC 308
218 GCGAGGTGCTGCACGGTCTGATTTGGGAGGACGACGAGGAGCGCGGGTGCACCAAGACC 277
309 ACCCGCGGCTGTGCTGTGTCGCCAGTAGCTCCGATTGGAATCATCCCGCAGGGTCAA 368
278 ACCCGCGGCTGTGCTGTGTCGCCAGTAGCTCCGATTGGAATCATCCCGCAGGGTCAA 337
369 CGGACTCGGTGTGTTACTCCACCGTGGGCACACGAGCGACGAACACCTTCGCGCTGCATA 428
338 CGGACTCGGTGTGTTACTCCACCGTGGGCACACGAGCGACGAACACCTTCGCGCTGCATA 397
429 TCGTGTGTTGGGACTCGTGTGCGCATGGATGTCTCAGTCACCAACAGGACACTCC 488
398 TCGTGTGTTGGGACTCGTGTGCGCATGGATGTCTCAGTCACCAACAGGACACTCC 457
489 TTCGCTACTCCGTGTCCCTGTGGGCTACCGGCTTCTACGGGACATCATCAAGACAGTG 548
458 TTCGCTACTCCGTGTCCCTGTGGGCTACCGGCTTCTACGGGACATCATCAAGACAGTG 517
549 AGAAGAAACGGTGTGGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTCT 608
|||||

Db 518 AGAAGAAACGGTGTGGGTCTTTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTCT 577
Qy 609 CCACCACTCTATGAA-GGACAGTGTCTCTCCCTCGACCAACACACACGGTGGATCT 667
Db 578 CCACCACTCTATGAAAGGGGACAGTGTCTCTCCCTCGACCAACACACGGTGGATCT 637
Qy 668 -CCAAGGATAGGAAGCCCTG-CGGGACAGATGCTTTG-TTTCAGGCAAGCAAGCAG 724
Db 638 CCNAGGATAGGAACCCCTGCCGGGACAGATGCTTTGTTGTCGGCAAGCAAGCAG 697
Qy 725 CAGCTGGAGGAGGACGAGAAAGC--ACTGTATGTTTGGAGCTGCGGAGACCTGG 782
Db 698 CAGCTGGAGGAGGACGAGAAACCACTGTATGTTTGGAACTGCGGAAAGCAGGC 757
Qy 783 AGGAGTGGCAAGTCTCT-CTGGGAGTCTCTGCCATCAATGCCACAAACATGCTCT-G 840
Db 758 AGGAGTGGCAATCTCTCTGGGGGAGTCTTCTGCCATCAATGCCCAACCTGTCTCTGG 817
Qy 841 TGCTGTGCGCGGAGCCCGGAGGCGCTCT---CCCGGCTGCCCACCTTGGGAGACGGGTC 897
Db 818 TGCTGTGCGCGGAGCCCGGAGGCGCTCTTCTCCCGGCTGCCCAATTTGGGAGAGGG 877
Qy 898 TTCT 901
Db 878 TCCT 881
RESULT 5
AK042077
LOCUS
DEFINITION
Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched
library, clone:A630056D11 product:DA59H18.2 (NOVEL PROTEIN SIMILAR
TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS)
(FRAGMENT) homolog [Homo sapiens], full insert sequence.
AK042077
VERSION AK042077.1 GI:26334912
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
92279253
PUBMED 10349636
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED 11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kiteuani,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsue,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED 11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
JOURNAL

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 6 (bases 1 to 1824)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Iehii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission
 JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge. Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/

FEATURES
 source

```
1..1824
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A630056D11"
/db_xref="taxon:10090"
/clone="A630056D11"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="3 days neonate"
```

85..1680

CDS
 /note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTR|Q9UE5, evidence: FASTA, 86.5%ID, 100%length, match=945) putative"

```
/codon_start=1
/protein_id="BAC31157.1"
/db_xref="GI:26334913"
/translation="MGANGAEPLHSLVWVKRRCAVSLPARALLRWRSPEPGPSA
PGADARSVLVSEITAVEEKDCERHASSGRWHKMNPAFTVHRVKRVRHHRWKRWAR
TFWSEDLQLHTQLTGLLESLSLRPKHLNVFINPFGGQGGKRIYKTVAPFLTL
ASITTEIITHANQAKETLYEINTDSYDGVICVGGDMGFSEVLHGVIGRTQQSAGID
PNHPRVLVPSLRIGIIPAGSTDCVSTGTNDATSAHLIIIGSLAIDSVSVHY
HNTLLRYSVLILGYFGDLIKDSEKHEWGLVRYDFSLGKATFLSHOYEGTILFLPA
QHTVGSPRDNKPCAGCFCVRSQKQLEEEBKALYLENAEEMEEQVTCGKFLAIN
ATNMSCAPRSPGSLSPFAHLGDSGSDILIRKCSRNFRLRHLRHTNQEDQDFTFV
EYRVKFFHFTSKHVEDNDNSKQEKQKFGKICKDRPSTCSASRSSNWDGVMHS
PAIEVRVHCVLRLFARGIEES"
```

ORIGIN

Query Match 64.9%; Score 635.8; DB 3; Length 1824;
 Best Local Similarity 79.6%; Pred. No. 2.9e-171;

| | Matches | 779; | Conservative | 0; | Mismatches | 167; | Indels | 33; | Gaps | 1; |
|----|---------|--------------------------------------------------------------|--------------|----|------------|------|--------|-----|------|----|
| Qy | 1 | ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAAGCAAGCAAGCGGAT | 60 | | | | | | | |
| Db | 471 | ACCGAAGCATTGCTGGTATTTATCAACCCCTTTTCGGAGGAAAGGTCAAGCGCAT | 530 | | | | | | | |
| Qy | 61 | ATATGAAAGAAATGGGACCACTGTTACCTTAGCCTTCATCAACCTGATCATCGG | 120 | | | | | | | |
| Db | 531 | CTATGAAAAAACAGTGGCGCTCTGTTTACCTTGGCTTCCATCACTACGGAGATCATC | 588 | | | | | | | |
| Qy | 121 | TAACAAATTTCTATGTTAACTATGTAGAAGTAATTAAGTGAACATGCTTAATCAAGG | 180 | | | | | | | |
| Db | 589 | -----ATTACAGAGCATGCCAACCAAGCAAGGA | 617 | | | | | | | |
| Qy | 181 | GACTCTGTATGAGATTAAATAGACGAGCATCGTCTGTGTGCGCGGAGATGG | 240 | | | | | | | |
| Db | 618 | GACTTTATACGAGATCAACAGACAGCATATGATGGCATCGTGTGGTAGGTGGGACGG | 677 | | | | | | | |
| Qy | 241 | TATGTTTACGAGAGTGTGCACCGTCTGATTTGGAGGACGAGAGGACGCGCGGGTCTGA | 300 | | | | | | | |
| Db | 678 | CATGTTTACGAGAGTGTGCATGGGGTGAATGGGAGGACGAGAGCGCGTGTGATCGA | 737 | | | | | | | |
| Qy | 301 | CCAGAACACCCCGGCGTGTGTGGTCCCGAGTAGCTCCGGATTGGAATCATTTCCGCG | 360 | | | | | | | |
| Db | 738 | CCCAATACCCCGAGCGGTGTGTGGTCCCGAGTAGCTCCAGTACCTTAGGATCGGCATAC | 797 | | | | | | | |
| Qy | 361 | AGGTCACAGGATGCGTGTGTATCTCACCGTGGGACCCAGCAGCAGCAAAACCTCGGC | 420 | | | | | | | |
| Db | 798 | AGGTCACAGATTTGTGTGTACTCAACAGTGGGACAAACGACGAGAGCATCGCG | 857 | | | | | | | |
| Qy | 421 | GCTGCATATGTTTGGGACCTCGTGGCATCGATGTCTCCTCAGTCCACCAACAG | 480 | | | | | | | |
| Db | 858 | TTTGACATCATTTATTTGGGACCTCACTGGCAATAGACGTGTCTCTGTGTGCACTACCA | 917 | | | | | | | |
| Qy | 481 | CACACTCTTCTGCTACTCCGTGCTCCGTGGGCTACGGCTTCTACGGGACATCATCAA | 540 | | | | | | | |
| Db | 918 | CACCTCTGCTGGTACTCGGTTTCTCTGCTGGGCTACGGTTTCTACGGGACTTAATCA | 977 | | | | | | | |
| Qy | 541 | GGACAGTGAGAAAGACGGTGGTGGTCTTTCGACAGATACGACTTTTCAGGTTTAAAGAC | 600 | | | | | | | |
| Db | 978 | GGACAGTGAAAGAAACGGTGGATGGGCTCGTCCGGTATGATTTCTCAGGTTTGAAGAC | 1037 | | | | | | | |
| Qy | 601 | CTTCTCTCCACACATGCTATGAAAGGACAGTGTCTTCTCTCCTTGCACAAACACCGGT | 660 | | | | | | | |
| Db | 1038 | CTTTCTCTCTCATCAGTACTATGAAAGGACACTGTCTCTTCTCCTCCAGCAGCACCGGT | 1097 | | | | | | | |
| Qy | 661 | GGGATCTTCAAGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTGACGGCAAGCAA | 720 | | | | | | | |
| Db | 1098 | GGGATCTTCAAGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTGACGGCAAGCAA | 1157 | | | | | | | |
| Qy | 721 | GCACAGCTGGAGGAGGAGCAAGAAAGCACTGTATGGTTTGAAGCTCGGAGGACGT | 780 | | | | | | | |
| Db | 1158 | GCAACAGCTGGAAGAAAGAGAGAAAGCCCTGTATGGCTTGAGAACGCGAGGAAT | 1217 | | | | | | | |
| Qy | 781 | GGAGAGTGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCAACAATGTCTGTG | 840 | | | | | | | |
| Db | 1218 | GGAAGAGTGGCAAGTGTGTGGAAAGTTTCTGGCCATCAATGCCAACAATGTCTGTG | 1277 | | | | | | | |
| Qy | 841 | TGCTTGTGCGCGAGGCCCTCCCGGCTGCCACCTTGGGAGACGGGTCTTC | 900 | | | | | | | |
| Db | 1278 | TGCTTGTGCTCGGAGCCCTCGGGGCCCTGTCCCCATTTGGCCATCTGGAGATGGGTCTTC | 1337 | | | | | | | |
| Qy | 901 | TGACCTCATCTCATCCGGAATGCTCCAGGTTTCAATTTTCTTGAGATTTCTCATCAGGCA | 960 | | | | | | | |
| Db | 1338 | TGACCTCATCTTATCCGGAAGTCTCCAGGTTTCAACTTCTTGAGATTTCTCATCCGCA | 1397 | | | | | | | |
| Qy | 961 | CACCAACAGGAGGACGAG | 979 | | | | | | | |
| Db | 1398 | CACCAACAGGAGGACGAG | 1416 | | | | | | | |

RESULT 6
 AK052269

| | | | | | |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|-----------|-----------------|
| LOCUS | AK052269 | 4248 bp | mRNA | linear | HTC 03-APR-2004 |
| DEFINITION | Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched library, clone:D330016D08 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence. | | | | |
| ACCESSION | AK052269 | 1 | GI:26342491 | | |
| VERSION | AK052269.1 | | | | |
| KEYWORDS | HTC; CAP trapper. | | | | |
| SOURCE | Mus musculus (house mouse) | | | | |
| ORGANISM | Mus musculus | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | |
| AUTHORS | 1 | Carninci, P., and Hayashizaki, Y. | | | |
| TITLE | High-efficiency full-length cDNA cloning | | | | |
| JOURNAL | Meth. Enzymol. | 303 | 19-44 | (1999) | |
| MEDLINE | 99279253 | | | | |
| PUBMED | 10349636 | | | | |
| REFERENCE | 2 | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | | | |
| AUTHORS | 3 | Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. | | | |
| TITLE | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer | | | | |
| JOURNAL | Genome Res. | 10 | (11) | 1757-1771 | (2000) |
| MEDLINE | 20530913 | | | | |
| PUBMED | 11076861 | | | | |
| REFERENCE | 4 | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. | | | |
| TITLE | Functional annotation of a full-length mouse cDNA collection | | | | |
| JOURNAL | Nature | 409 | 685-690 | (2001) | |
| REFERENCE | 5 | The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. | | | |
| TITLE | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs | | | | |
| JOURNAL | Nature | 420 | 563-573 | (2002) | |
| REFERENCE | 6 | (bases 1 to 4248) | | | |
| AUTHORS | Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216] | | | | |
| COMMENT | cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken | | | | |

| | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/. | |
| FEATURES | location/Qualifiers |
| source | 1. 4248
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:D330016D08"
/db_xref="taxon:10090"
/clone="D330016D08"
/tissue type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
260.11855
/note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTR{Q9UG55, evidence: FASTY, 86.5%ID, 100%length, match=945) putative"
/codon_start=1
/protein_id="BAC34908.1"
/db_xref="GI:26342492"
/translations="MGAMAAEPLHSLVWVKRRRCASVLEPARALLRWMRSPGPGSPA
PGDARSQVLVSEIIIAVEEKDDCEKHASSGRWHKMPAFTVHRVKRVHRHKKWAVR
TFWSADQLCHLWQTLRGLLESLSRPHLLVFINPFGKGQGRKRIEYKTVAQPLFTL
ASITTEIIIEHANQAEILYEINTSDYICVCGDGMFSEVLHGKIVTQPSAGID
PNHPRLVPSLIRIGIIPAGSDVCYCTVGTNDABTSALHIIIGDSLAIIDVSVHY
HNTLLRVSVLLGYGFYDILIKDSEKRWMLGVYDFSLGTLFLSHQYVEGTLSPFA
ATWMSCAPRSPGSLSPFAHLPGSSDLILIRKCSRFNLFRLIRHTNOEDOFDTFV
EVRVYKPHFTSKHVEYEDNDSKEQKFKICDRPSTCSASRSSNWCGEVMSHS
PAIEVRVHCOLVRLFARGIEES" |
| ORIGIN | Query Match 64.9%; Score 635.8; DB 3; Length 4248;
Best Local Similarity 79.6%; Pred. No. 3.7e-171;
Matches 779; Conservative 0; Mismatches 167; Indels 33; Gaps 1;

QY 1 ACCAAAGCATTTACTGCTATTATTAACAACCGTTTGAGGAGAAAAGCAAGCGGAT 60

DB 646 ACCGAAGCATTGCTGCTATTTCATCAACCTTTCCGAGGGAAAGTCAGGCAAGCCAT 705

QY 61 ATATGAAGAAAAGTGCACACCTGTTTCACCTTAGCTCCATCACCACCTGACATCATCG 120

DB 706 CTATGAAAAAACAGTGGCGCTCTCTGTTTACCTTGGCTTCATCACTACCGAGATCATC-- 763

QY 121 TAACAAATTCATGTTAACTATGTAGAAGTAATTTACTGAACATGCTAATCAGGCCAAGGA 180

DB 764 -----ATTACAGACATGCCAACCAAGCCCAAGGA 792

QY 181 GACTCTGTATGAGATTAACTAGACAAATAACGACGGCATCGTCTGTCTGGCGGAGATGG 240

DB 793 GACTTTATACGAGATCAACACACACAGCTATGATGGCATCGTCTGTGTAGTGGGACGG 852

QY 241 TATGTTACGAGAGTGTGTCACGGTCTGATTGGGAGAGACGACAGAGAGCGCCGGGTCGA 300

DB 853 CATGTTTCAGCGAGGTGCTGCTGGGTGGTGGTGGGAGACGACAGACAGAGCGCTGGTATCGA 912

QY 301 CCAGAACACCCCGGGCTGTGCTGGTCCCGTAGCTCCCGATTGGGAATCATTCCTCCGC 360

DB 913 CCCCAATCACCCCGAGCCGTGCTGGTGCCAGTACCTCAGATCGGATCATATCCCGC 972

QY 361 AGGGTCAACGGACTCGCTGTGTTACTCCACCGTGGGACCCAGCGACGAGAAACCTTCGCG 420

DB 973 AGGGTCCACAGATTGTGTGTGTACTCAACAGTGGGCACAAACGACGACAGACATCGCG 1032

QY 421 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCTCTAGTCCACCAACAG 480

DB 1033 TTTCACATCATATTGGGGACTCATTGGCAATAGACGTGTCTCTGTGCATTCACATAA 1092

DB 1033 TTTCACATCATATTGGGGACTCATTGGCAATAGACGTGTCTCTGTGCATTCACATAA 1092
 |

| | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| FEATURES | source |
| 1..4248 | Location/Qualifiers |
| /organism="Mus musculus" | |
| /mol_type="mRNA" | |
| /strain="C57BL/6J" | |
| /db_xref="PANTOM DB:D330016D08" | |
| /db_xref="taxon:10090" | |
| /clone="D330016D08" | |
| /tissue_type="heart" | |
| /clone_lib="RIKEN full-length enriched mouse cDNA library" | |
| /dev_stage="13 days embryo" | |
| 260..1855 | |
| /note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTR Q9UG5, evidence: FASTY, 86.5%ID, 100%length, match=945) putative" | |
| /codon_start=1 | |
| /protein_id="BAC34908.1" | |
| /db_xref="GI:26342492" | |
| /translation="MGAMGAEP LHSVLVWRRCVSLPEPARALLRWPRSPGPSA PGADARSVLVSEIIAIVEEKDDCKEHAASSGRHWKMNPPAFVHRVVRHHRWKARV TFSADDEQLCHLWLQTLRLLESITDSYGIIVGCGDGMFSEVLHGYSQTSAGID HPNPRVLPVSTRIIGIIPAGSTDVCYVTGNTDAETSAHIIIGSLAIDVSSVHY HNTLRYSVLLGCVGFVGLIKDSEKRWMLVRYDFSLGKTLFSLHQYVGTLSFLPA OHTVGSPPRONKCRACFCVCRSQKOLEEKKALXGLENAEMESQVTCGRFLAIN ATNMSACAPRSPGGLSPFAHLGDGSSDLILIRKCSRPNFLRLIRHTNQEDQDFTFV EVTRVKRFPSTSKHVEYENDSKQKQKFKIKDKRPSCTCSASRSSMNCDEGVMS PAIEVVRHVCQLVRLPARGTEES" | |
| ORIGIN | |
| Query Match 64.9%; Score 635.8; DB 3; Length 4248; | |
| Best Local Similarity 79.6%; Pred. No. 3.7e-171; | |
| Matches 779; Conservative 0; Mismatches 167; Indels 33; Gaps 1; | |
| Qy 1 ACCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAAAGGCAAGCGAAGCGGAT 60 | |
| | |
| 646 ACCGAGCACTTCTGCTGTTATTCATCAACCTTTTCGGAGGAAAGGTCAGGCGAAGCGCAT 705 | |
| | |
| Qy 61 ATATGAAAGAAAAGTGGCACCCTGTTACCTTAGCCTCCATCACCACTCACTATCATCGG 120 | |
| | |
| 706 CTATGAAAAAACAGTGGCGCTCTGTTTACCTTTGGCTTCCATCACTACGAGAGATCATC-- 763 | |
| | |
| Qy 121 TAACAAATTCATGTTAACTATGTAGAAGTAATTACTGAACATGCTAATCAGGCCAAGGA 180 | |
| | |
| 764 -----ATTACAGACATGTCACCAAGCCAAGGA 792 | |
| | |
| Qy 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTGCGCGGAGATGG 240 | |
| | |
| 793 GACTTTATACGAGATCAACACACACAGCTATATGATGATGATGATGATGATGATGATGAT 852 | |
| | |
| Qy 241 TATGTTTCAGGAGGTGTCGACCGTCTGATGGAGAGACCCAGAGAGCGCCGCGGTTCGA 300 | |
| | |
| 853 CATGTTTCAGGAGGTGTCGATGGGTGATGTTGGAGAGACGACGACGACGCGCTGATTCGA 912 | |
| | |
| Qy 301 CCAGAACCCCGGCTGTGTCGCCAGTAGCTCCGATCGGATCGGAATCATTCCTCCGC 360 | |
| | |
| 913 CCCCAATACCCCGAGCCGTGCTGTGCCAGTACCTCAGATCGGATCATACCCGC 972 | |
| | |
| Qy 361 AGGCTCAACGAGCTGCGTGTGTACTCCACCGTGGGACCCAGCGACCGCAGAAAACCTCGGC 420 | |
| | |
| 973 AGGCTCCACAGATGTTGTGTGTACTCAACAGTGGGACAAACGACGACGAGACATCGGC 1032 | |
| | |
| Qy 421 GCTGCATATCGTTGTTGGGACCTCGTGGCCATGGATGTGTCCTCAGTCCACCAACACAG 480 | |
| | |
| 1033 TTTGCAATCATATTATTTGGGACCTCACTGGCAATAGACGTGTCTCTGTGCACTACCATAA 1092 | |
| | |


```
QY 481 CACACTCTTCCTACTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAA 540
Db 1093 CACGCTGCTCGGTACTCGGTTTCTGCTGGGCTACGGTTTCTACGGGGACTTAATCAA 1152
QY 541 GGACAGTGAGAAACGGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGAC 600
Db 1153 GGACAGTGAAAGAAACGGTGATGGGCTCTGCGGTATGATTTCTCAGGTTGAAGAC 1212
QY 601 CTTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGACACACACACGGT 660
Db 1213 CTTTCTCTCATAGTACTATGAAGGACACTGTCTTCTCTCCAGACACACACGGT 1272
QY 661 GGGATCTCAAGGGATAGAAAGCCCTGCGGGGACAGATGCTTTGTTGGAGCAAAAGCAA 720
Db 1273 GGGATCTCCAGGGCAATAAACCCTGCGGGGCTGGTCTGCTGTCGAGGACAGCAA 1332
QY 721 GGACAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAGCTGCGGAGGAGT 780
Db 1333 GCAACAGCTGGAAGAAAGAAAGAAAGCCCTGTATGGCTTGGAGAACGCGGAGGAAAT 1392
QY 781 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTCTGGCCATCAATGCCACAAACATGTCCTG 840
Db 1393 GGAAGTGGCAAGTGACATGTGGGAAGTTCTGGCCATCATGCCACACATGTCCTG 1452
QY 841 TGCTGTGCGCGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGAGGGTCTTC 900
Db 1453 TGCTGTGCTCGAGAGCCCTGCGGGGCTGTCCCAATTTGCCATCTGGGAGATGGGTCTTC 1512
QY 901 TGACCTCATCTTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTCATCAGGCA 960
Db 1513 TGACCTCATCTTATCCGGAAGTGTCCAGGTTCAACTTCTGAGATTCCTCATCCGCA 1572
QY 961 CACCAACAGCAGGACGACAG 979
Db 1573 CACGACACGAGGACGACAG 1591
```

```
RESULT 7
LOCUS CD655311 820 bp mRNA linear EST 18-JUN-2003
DEFINITION AGENCOURT_14552675 NIA Human HI Embryonic Stem Cell cDNA Library
(Long) Homo sapiens cDNA clone IMAGE:30426593 5', mRNA sequence.
ACCESSION CD655311
VERSION CD655311.1 GI:31895467
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
Contact: Daniela S. Gerhard, Ph.D.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/ILLNL at:
http://image.llnl.gov
Plate: NDAM512 row: k column: 18
High quality sequence stop: 673.
Location/Qualifiers
1..820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```

```
/clone="IMAGE:30426593"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human HI Embryonic Stem Cell cDNA Library
(Long)"
/notes="vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001). [PMID: 1154199]) from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from WiCell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MEF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, SSEA-1,
TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days
after plating), the ES cells from 4 X 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(invitrogen/GIBCO) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with Trizol Reagent
from invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID:1154199]) Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen: 5'-
5'-pGACTAGTCTAGATCGGAGCGGCCCTTTTITTTT-3'] from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lona-linker LL-Sal4, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
pCMV-SPORT6 plasmid vector. The average insert size is
about 3.6kb."
```

ORIGIN

```
Query Match 60.6%; Score 593.4; DB 6; Length 820;
Best Local Similarity 99.8%; Pred. No. 3.6e-159;
Matches 594; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 385 CTCACCCGTGGGACCGACGACGAGAACCTCGGCGCTGCATATCGTTGTGGGACTC 444
Db 11 CTCACCCGTGGGACCGACGACGAGAACCTCGGCGCTGCATATCGTTGTGGGACTC 70
QY 445 GCTGCCATGATGTGTCTCAGTCCACCAACAGACACACTCTCTGCTACTCGGTGC 504
Db 71 GCTGCCATGATGTGTCTCAGTCCACCAACAGACACACTCTCTGCTACTCGGTGC 130
QY 505 CCTGCTGGGTACGGCTTCTACGGGGACATCATCAAGGACAGTGGAGAGAACGGTGT 564
Db 131 CCTGCTGGGTACGGCTTCTACGGGGACATCATCAAGGACAGTGGAGAGAACGGTGT 190
QY 565 GGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCCACTGCTATGA 624
Db 191 GGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCCACTGCTATGA 250
QY 625 AGGACAGTGTCTTCTCTGTCACACACACCGGTGGGATCTCCAGGGATAGGAAGCC 684
Db 251 AGGACAGTGTCTTCTCTGTCACACACACCGGTGGGATCTCCAGGGATAGGAAGCC 310
QY 685 CTGCGGGCAGGATGCTTGTTCAGGCAAGCAACAGCAGCTGGAGAGGAGCAGAA 744
Db 311 CTGCGGGCAGGATGCTTGTTCAGGCAAGCAACAGCAGCTGGAGAGGAGCAGAA 370
QY 745 GAAAGCACTGTATGCTTTGGAGGAGCTGGAGGAGCTGGAGGAGTCTGTGG 804
Db 371 GAAAGCACTGTATGCTTTGGAGGAGCTGGAGGAGCTGGAGGAGTCTGTGG 430
QY 805 GAAGTTTCTGCCATCAATGCCACAAACATGCTGTGTTGTCGCCGAGCCCCAGGG 864
|||||
```

```
Db 431 GAAGTTTCGGCCATCAATGCAAAACATGTCCTGTGCTTGTGCGCGGAGCCCGACGGG 490
Qy 865 CCTCTCCCGCGCTGCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAATG 924
Db 491 CCTCTCCCGCGCTGCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAATG 550
Qy 925 CTCGAGGTCAATTTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACACG 979
Db 551 CTCGAGGTCAATTTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACACG 605

RESULT 8
BQ063738
LOCUS BQ063738
DEFINITION AGENCOURT_6873251 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925382
5', mRNA sequence.
ACCESSION BQ063738
VERSION BQ063738.1 GI:19891754
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1078)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2096 row: p column: 23
High quality sequence stop: 640.
FEATURES
Location/Qualifiers
1..1078
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5925382"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 59.5%; Score 582.4; DB 5; Length 1078;
Best Local Similarity 92.6%; Pred. No. 5.8e-156;
Matches 659; Conservative 0; Mismatches 16; Indels 37; Gaps 3;
Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAGACAAAGCGCGAT 60
Db 120 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAGACAAAGCGCGAT 179
Qy 61 ATATGAAGAAAGTGGCACCACCTGTTTCACTTAGCTCCATCCACTGACATCATCGG 120
Db 180 ATATGAAGAAAGTGGCACCACCTGTTTCACTTAGCTCCATCCACTGACATCATCG- 238
Qy 121 TAACAAAATTCATGTTAACTATGTAGAAGTAAATTACTGAACATGCTATCAGGCAAGGA 180
Db 239 -----TTACTGAACATGCTATCAGGCAAGGA 266
```

```
Qy 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGCATCGTCTGTGTCGGCGAGATGG 240
Db 267 GACTCTGTATGAGATTAAACATAGACAAATACGACGCATCGTCTGTGTCGGCGAGATGG 326
Qy 241 TATGTTTCAGGAGGTGCTGACCGGTCTGATTTGGAGAGACGACAGAGAGCCCGGGGTCTGA 300
Db 327 TATGTTTCAGGAGGTGCTGACCGGTCTGATTTGGAGAGACGACAGAGAGCCCGGGGTCTGA 386
Qy 301 CCAGAACCCCGCGGCTGTGTCGTCCTCCAGTACGCTCCGATTTGGATTCATTCCTCCG 360
Db 387 CCAGAACCCCGCGGCTGTGTCGTCCTCCAGTACGCTCCGATTTGGATTCATTCCTCCG 446
Qy 361 AGGTCACACCGACTGCGTGTGTTACTCCACCGTGGGCACGACGACGACCAAACTCCGGC 420
Db 447 AGGTCACACCGACTGCGTGTGTTACTCCACCGTGGGCACGACGACGACCAAACTCCGGC 506
Qy 421 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTAGTCCACCAACAG 480
Db 507 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTAGTCCACCAACAG 566
Qy 481 CACACTCTTCGCTACTCCGCTGTCCTGCTGGGCTTCTACGGGACATCATCAA 540
Db 567 CACACTCTTCGCTACTCCGCTGTCCTGCTGGGCTTCTACGGGACATCATCAA 626
Qy 541 GGCAGTGAAGAAGAACCGGTGGTGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGAC 600
Db 627 GGCAGTGAAGAAGAACCGGTGGTGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGAC 686
Qy 601 CTTCTCTCCACACACTGCTATGAAGGACAGTGTCTTCTCTCCCTGTCACACACAGG- 659
Db 687 CTTCTCTCCACACACTGCTATGAAGGACAGGTCCTTCTCTCTGTCACACACAGG 746
Qy 660 -TGGGATCTCCAAAGGA--TAGGAAGCCCTGCCGGCAGGATGCTTTGTTTG 708
Db 747 TGGGGATCTCCAGGGAATAAGGAAGCCCTGCCGGCAGGATGCTTTG 798

RESULT 9
BQ224560
LOCUS BQ224560
DEFINITION BQ224560 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA
clone DAT01067, mRNA sequence.
ACCESSION BQ224560
VERSION BQ224560.1 GI:52097465
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 581)
REFERENCE Suzuki,Y., Yamashita,R., Hirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
JOURNAL Contact: Yutaka Suzuki
COMMENT Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezuk@iims.u-tokyo.ac.jp.
FEATURES
Location/Qualifiers
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DAT01067"
/cell_type="lymphocyte"
/cell_line="Daudi"
/clone_lib="Sugano cDNA library, lymphocyte Daudi"
/note="Burkitt's lymphoma"
ORIGIN
```

| | | | | | |
|-----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|-----------|-------------|-----------------|
| Query Match | 59.2%; | Score 579.4; | DB 5; | Length 581; | |
| Best Local Similarity | 99.8%; | Pred. No. 3.5e-155; | | | |
| Matches 580; | Conservative 0; | Mismatches 1; | Indels 0; | Gaps 0; | |
| QY | 195 | TTAAATAGACAAATACGACGGCATCGTGTGTGGCGGAGATGATGTTTCAGCGAGG | 254 | | |
| Db | 1 | TTAAATAGACAAATACGACGGCATCGTGTGTGGCGGAGATGATGTTTCAGCGAGG | 60 | | |
| QY | 255 | TGCTGCACGGTCTGATTTGGGAGGACGACAGAGAGCGCGGGGTGCACCAAGAACCCACCC | 314 | | |
| Db | 61 | TGCTGCACGGTCTGATTTGGGAGGACGACAGAGAGCGCGGGGTGCACCAAGAACCCACCC | 120 | | |
| QY | 315 | GGGCTGTGTCGTCCCGTACGCTCCGATTTGGATCATTTCCCGAGGTTCAACGGACT | 374 | | |
| Db | 121 | GGGCTGTGTCGTCCCGTACGCTCCGATTTGGATCATTTCCCGAGGTTCAACGGACT | 180 | | |
| QY | 375 | GGGTGTGTTACTCCACCGTGGGACACGACGACGAGAACTCGCGGCTGCATATCGTTG | 434 | | |
| Db | 181 | GGGTGTGTTACTCCACCGTGGGACACGACGACGAGAACTCGCGGCTGCATATCGTTG | 240 | | |
| QY | 435 | TTGGGAGCTCGTGGCCATGGATGTCTCTAGTCCACCAACACAGCACACTCTTCGCT | 494 | | |
| Db | 241 | TTGGGAGCTCGTGGCCATGGATGTCTCTAGTCCACCAACACAGCACACTCTTCGCT | 300 | | |
| QY | 495 | ACTCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATCAAGACAGTGAAGA | 554 | | |
| Db | 301 | ACTCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATCAAGACAGTGAAGA | 360 | | |
| QY | 555 | AACGGTGTGGTGTTCGCAGATGACGCTTTTCAGGTTTAAAGACCTTCCTCTCCAC | 614 | | |
| Db | 361 | AACGGTGTGGTGTTCGCAGATGACGCTTTTCAGGTTTAAAGACCTTCCTCTCCAC | 420 | | |
| QY | 615 | ACTGCTATGAAGGACAGTGTCTCTCTCCGCAACACACGCTGGGATCTCAAGGG | 674 | | |
| Db | 421 | ACTGCTATGAAGGACAGTGTCTCTCTCCGCAACACACGCTGGGATCTCAAGGG | 480 | | |
| QY | 675 | ATAGAAGCCCTCGCGGAGAGATGCTTTGTTGAGGCAAGCAGCAGCTGGAGG | 734 | | |
| Db | 481 | ATAGAAGCCCTCGCGGAGAGATGCTTTGTTGAGGCAAGCAGCAGCTGGAGG | 540 | | |
| QY | 735 | AGGACGAGAAGAAGCACTGTATGTTGGAAGCTCGCGAG | 775 | | |
| Db | 541 | AGGACGAGAAGAAGCACTGTATGTTGGAAGCTCGCGAG | 581 | | |
| RESULT 10 | | | | | |
| CB246749 | | | | | |
| LOCUS | CB246749 | 758 bp | mRNA | linear | EST 09-JUL-2003 |
| DEFINITION | UI-M-F10-cdx-b-10-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone | | | | |
| ACCESSION | IMAGE:6835595 5', mRNA sequence. | | | | |
| VERSION | CB246749 | | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | Mus musculus (house mouse) | | | | |
| ORGANISM | Mus musculus | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| AUTHORS | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | |
| JOURNAL | 1. (bases 1 to 758) | | | | |
| COMMENT | NIH-MGC http://mgc.nci.nih.gov/ .
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5. | | | | |

| | | | | | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|--------------------------------------------------------------|-----------|-------------|--|
| Location/Qualifiers | | | | | |
| 1..758 | | | | | |
| /organism="Mus musculus" | | | | | |
| /mol_type="mRNA" | | | | | |
| /strain="CS7BL/6" | | | | | |
| /db_xref="taxon:10090" | | | | | |
| /clone="IMAGE:6835595" | | | | | |
| /tissue_type="whole brain" | | | | | |
| /dev_stage="embryo 12.5dpc" | | | | | |
| /lab_host="DH10B (T1 phage resistant)" | | | | | |
| /clone_lib="NIH_BMAP_F10" | | | | | |
| /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator." | | | | | |
| ORIGIN | | | | | |
| Query Match | 54.9%; | Score 537.4; | DB 6; | Length 758; | |
| Best Local Similarity | 82.7%; | Pred. No. 4.7e-143; | | | |
| Matches 613; | Conservative 0; | Mismatches 128; | Indels 0; | Gaps 0; | |
| QY | 239 | GGTATGTTTCAGCGAGGTCTGCACGGTCTGATTTGGGAGGACGACGAGGAGCCCGGGTTC | 298 | | |
| Db | 1 | GGCATGTTTCAGCGAGGTCTGCATGGGGTGTATGGGAGGACGACGAGCGCTGGTATC | 60 | | |
| QY | 299 | GACCAAGAACACCCCGGGGTGTGTCGTCCCGTAGCTCCGAGTGGAAATCATTTCCC | 358 | | |
| Db | 61 | GACCCCAATACCCCGGAGCGGTGCTGTTGCCAGTACCCCTCAGATCGGCATCATACCC | 120 | | |
| QY | 359 | GCAGGCTCAACGGACTCGGTGTGTACTCCACCGTGGGACGACGAGGAGAAACCTCG | 418 | | |
| Db | 121 | GCAGGCTCCACAGATTGTGTGTACTCAACAGTGGGACAAACGACGAGAGACATCG | 180 | | |
| QY | 419 | CGCTGTCATATCGTTGTTGGGAGCTCGCTGCGCCATGGATGTCTCTCAGTCCACCAAC | 478 | | |
| Db | 181 | GCTTTGCACATCATTTATTTGGGAGCTCACTGGCAATAGACGTCTCTCTGTGCATACCAT | 240 | | |
| QY | 479 | AGCACATCTCTTCGCTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGGACATCATC | 538 | | |
| Db | 241 | AACACGCTGCTGCGGTACTCGGTTTCTCTGCTGGGCTACGGTTTCTACGGGACTTAATC | 300 | | |
| QY | 539 | AAGGACAGTGAGAGAAACGGTGGTGTGCTTCCAGATACGACTTTTCAGGTTTAAAG | 598 | | |
| Db | 301 | AAGGACAGTGAAGAAAGAAACGGTGGATGGCTCTGTCGGTATGATTTCTCAGGGTTGAAG | 360 | | |
| QY | 599 | ACCTTCTCTCTCCACCACTGCTATGAAGGAGCAGTGTCTCTCTCTCCCTGCACCAACACG | 658 | | |
| Db | 361 | ACCTTCTCTCTCATCAGTACTATGAAGGAGACGTCTCTCTCTCCAGCAGCAGCAGCAG | 420 | | |
| QY | 659 | GTGGATCTCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGAGGCAAGC | 718 | | |
| Db | 421 | GTGGATCTCCACGGGACATAAACCCTGCGGGCTGGGTGCTTGTGTGAGGAGCAGCAGC | 480 | | |
| QY | 719 | AAGCAGCAGCTGGAGGAGCAGAGAAAGCACTGTATGTTTGGAGCTCCGAGGAC | 778 | | |
| Db | 481 | AAGCAACAGCTGGAGAGAGAGAAAGAACCCCTGTATGGCTGGAGAAACGCGGAGAA | 540 | | |
| QY | 779 | GTGGAGGAGTGGCAAGTCTGTGTGGGAAAGTTTCTGGCCATCAATGCCACAAACATGTCC | 838 | | |
| Db | 541 | ATGGAGAGTGGCAAGTGCATGTGGGAAAGTTCTTGGCCATCAATGCCACCAACATGTCC | 600 | | |
| QY | 839 | TGTGCTTGTCCGCGAGCCCGGCGGCTCTTCCCGGCTGCCCACTTGGGAGACGGGTCT | 898 | | |


```

Db      123  TGTGGGACTCGCTGGCCATGATGTGCTCCTAGTCCACCAACAGCACACTCCTTCG 182
Qy      493  CTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGAGAA 552
Db      183  CTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGAGAA 242
Qy      553  GAAACGGTGGTGGGCTGTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCCA 612
Db      243  GAAACGGTGGTGGGCTGTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCCA 302
Qy      613  CCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACAGTGGGATCTCCAG 672
Db      303  CCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACAGTGGGATCTCCAG 362
Qy      673  GGATAGGAAGCCCTCGCGGCGAGGATGCTTTTTCAGGCGAAAGCAAGCAGCTGGA 732
Db      363  GGATAGGAAGCCCTCGCGGCGAGGATGCTTTTTCAGGCGAAAGCAAGCAGCTGGA 422
Qy      733  GGAGAGCAGAAAGACACTGTATGTTTGAAGCTGCGGAGGACGTGGAGGATGGCA 792
Db      423  GGAGAGCAGAAAGACACTGTATGTTTGAAGCTGCGGAGGACGTGGAGGATGGCA 482
Qy      793  AGTGTCTGTGGGAAAGTTTCTCGCCATCAATGCCAACAACATGTCCTGTGCTGTGCGC 851
Db      483  AGTGTCTGTGGGAAAGTTTCTCGCCATCAATGCCAACAACATGTCCTGTGCTGTGCGC 541

RESULT 13
BP309990 Sugano cDNA library, brain Homo sapiens EST 17-SEP-2004
LOCUS NR01216, mRNA sequence.
ACCESSION BP309990
VERSION BP309990.1 GI:52238965
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini, Hominidae; Homo.
TITLE Suzuki,Y., Yanashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1. 584
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NR01216"
/tissue_type="brain"
/clone_lib="Sugano cDNA library, brain"

ORIGIN
Query Match 53.9%; Score 527.4; DB 5; Length 584;
Best Local Similarity 94.2%; Pred. No. 3.3e-140;
Matches 582; Conservative 0; Mismatches 2; Indels 34; Gaps 2;
Qy 29 CCGTTGGAGGAAAGGACAGCGGATATGAAAGAAAGTGGCACCACCTGTTTC 88
Db 1 CCGTTGGAGGAAAGGAAAGGACAGCGGATATGAAAGAAAGTGGCACCACCTGTTTC 59
Qy 89 ACCTTAGCTCCATCACCACCTCAGATCATCGGTAAACAAATCTATGTTAACTATAGAA 148
Db 60 ACCTTAGCTCCATCACCACCTCAGATCATCG----- 90
```

```

Qy      149  GTAATTCTGAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAA 208
Db      91  ---TTACTGAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAA 146
Qy      209  TACGACGCGATCCTGCTGTGTCGGCGGAGATGGTATGTTTTCAGGAGGTCCTGCACGGTCTG 268
Db      147  TACGACGCGATCCTGCTGTGTCGGCGGAGATGGTATGTTTTCAGGAGGTCCTGCACGGTCTG 206
Qy      269  ATTGGGAGGACGACAGAGGAGCGCCGGGTCGACCAAGAACCCCGCGGCTGTGCTGGTTC 328
Db      207  ATTGGGAGGACGACANAGGAGCGCCGGGTCGACCAAGAACCCCGCGGCTGTGCTGGTTC 266
Qy      329  CCCAGTACCTCCCGGATTTGGAATCATTTCCCGCAGGCTCAACGAGTCTGTGTTACTCC 388
Db      267  CCCAGTACCTCCCGGATTTGGAATCATTTCCCGCAGGCTCAACGAGTCTGTGTTACTCC 326
Qy      389  ACCGTGGGACACAGCGGACGAGAAACCTTCGCGCTGCATATCGTTGTTGGGACTCGCTG 448
Db      327  ACCGTGGGACACAGCGGACGAGAAACCTTCGCGCTGCATATCGTTGTTGGGACTCGCTG 386
Qy      449  GCCATGATGTCTCTCAGTCCACCAACAGCAGCACCTCTTTCGCTACTCCGTCCTCCCTG 508
Db      387  GCCATGATGTCTCTCAGTCCACCAACAGCAGCACCTCTTTCGCTACTCCGTCCTCCCTG 446
Qy      509  CTGGCTACGGCTTCTACGGGACATCATCAAGAGCAGTGAAGAAACGTTGGTGGGT 568
Db      447  CTGGCTACGGCTTCTACGGGACATCATCAAGAGCAGTGAAGAAACGTTGGTGGGT 506
Qy      569  CTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTCTCCACCACTGCTATCAAGGG 628
Db      507  CTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTCTCCACCACTGCTATCAAGGG 566
Qy      629  ACAGTGTCTCTCTCCCTCCCT 646
Db      567  ACAGTGTCTCTCTCCCTCCCT 584

RESULT 14
CFI38275 573 bp mRNA linear EST 09-SEP-2003
LOCUS UI-HF-BNO-anz-g-11-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone
DEFINITION IMAGE:3094461 5', mRNA sequence.
ACCESSION CFI38275
VERSION CFI38275.1 GI:33253719
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalizatiion and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEERF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Louis Staudt
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pyx-5.
FEATURES
source
Location/Qualifiers
1. 573
/organism="Homo sapiens"
```

```

/molecule="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3094461"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LTI)"
/clone_lib="NIH_MGC_50"
/notes="Vector: pT7T3-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN
Query Match 53.2%; Score 520.8; DB 7; Length 573;
Best Local Similarity 93.9%; Pred. No. 2.6e-138;
Matches 565; Conservative 0; Mismatches 4; Indels 33; Gaps 1;

Qy 53 AAGCGGATATATGAAGAAAGTGGCCACCTGTTACCTTAGCTTCATCACCACCTGAC 112
Db 5 AGGCGGATATATGAAGAAAGTGGCCACCTGTTACCTTAGCTTCATCACCACCTGAC 64
Qy 113 ATCATCGGTAAACAAATCTTATGTAACTATGTAAGATTAATCTGAACATGCTAAATCAG 172
Db 65 ATCATCG-----TTACTGAACATGCTAAATCAG 91
Qy 173 GCCAAGGAGACTCTGTATGAGATTAACATAGACAAATACGCGCATCGTGTGTGGC 232
Db 92 GCCAAGGAGACTCTGTATGAGATTAACATAGACAAATACGCGCATCGTGTGTGGC 151
Qy 233 GGAGATGTTATGTTTCAGCGAGTGTCTGACCGTCTGATTTGGAGAGACGACGAGGAGGCC 292
Db 152 GGAGATGTTATGTTTCAGCGAGTGTCTGACCGTCTGATTTGGAGAGACGACGAGGAGGCC 211
Qy 293 GGGTTCGACCAAGAACCCCGGGCTGTCTGTGTCCTAGTACGCTCCGATTTGGAATC 352
Db 212 GGGTTCGACCAAGAACCCCGGGCTGTCTGTGTCCTAGTACGCTCCGATTTGGAATC 271
Qy 353 ATTCGCGAGGTCACCGAGTGTGTTACTCCACCGTGGGACGACGAGCGCAGAA 412
Db 272 ATTCGCGAGGTCACCGAGTGTGTTACTCCACCGTGGGACGACGAGCGCAGAA 331
Qy 413 ACCTCGGGCTGCATATCTGTTGGGACTCGCTGGCCATGATGTCTCCTCAGTCCAC 472
Db 332 ACCTCGGGCTGCATATCTGTTGGGACTCGCTGGNCATGATGTCTCCTCAGTCCAC 391
Qy 473 CACAACAGCACATCTCTTCGCTACTCTCGTGTCCCTGTGGGCTACGGCTTCTACGGGAC 532
Db 392 CACAACAGCACATCTCTTCGCTACTCTCGTGTCTGTCGCTGGGCTTCTACGGGAC 451
Qy 533 ATCATCAAGGACAGTGAAGAACCGTGTGGTCTTGGCTTTCGCAGATACGACTTTTCAGGT 592
Db 452 ATCATCAAGGACAGTGAAGAACCGTGTGGTCTTGGCTTTCGCAGATACGACTTTTCAGGT 511
Qy 593 TTAAGACCTTCTCTCCACACCTGCTATGAAGGACAGTGTCTCTCTCCCTCCCTGCACAA 652
Db 512 TTAAGACCTTCTCTCCACACCTGCTATGAAGGACAGTGTCTCTCTCTCTCTCTGACAA 571
Qy 653 CA 654
Db 572 CA 573
```

```

RESULT 15
CF138634
LOCUS
DEFINITION
IMAGE:3094594 5', mRNA sequence.
ACCESSION
CF138634
VERSION
CF138634.1 GI:33254078
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
CF138634
UI-HF-BNO-aoc-b-06-0-UI_r1 NIH_MGC_50 Homo sapiens cDNA clone
IMAGE:3094594 5', mRNA sequence.
EST.
Homo sapiens (human)
```

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 550)

AUTHORS

Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

JOURNAL

discovery

MEDLINE

Genome Res. 6 (9), 791-806 (1996)

PUBMED

8889548

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Louis Staudt

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

http://genome.uiowa.edu/distribution/humanfl.html

Seq primer: pYX-5.

Location/Qualifiers

1..550

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3094594"

/tissue_type="lymph"

/cell_type="germinal center B cells"

/cell_line="MGC85"

/lab_host="DH10B (LTI)"

/clone_lib="NIH_MGC_50"

/notes="Vector: pT7T3-Pac; Site 1: NotI; Site 2: Eco RI;

Constructed from size fractionated cytoplasmic mRNA

(3.5-4.4kb). Directionally cloned. Cells provided by

Louis M. Staudt, Ph.D. Library preparation by Maria de

Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

Query Match 52.9%; Score 518.2; DB 7; Length 550;

Best Local Similarity 99.2%; Pred. No. 1.4e-137;

Matches 520; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 456 ATGTGTCTCAGTCCACCAACAGCACCTCTTCCTACTCGTCTCGTCTCGGCT 515

Db 2 ACGAGGCTCAGTCCACCAACAGCACCTCTTCCTACTCGTCTCGTCTCGGCT 61

Qy 516 ACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGTGGGTCTTGCCA 575

Db 62 ACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGTGGGTCTTGCCA 121

Qy 576 GATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGTCTATGAAGGACAGTGT 635

Db 122 GATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGTCTATGAAGGACAGTGT 181

Qy 636 CTTTCCTCCCTGCACACACACGCTGGGATCTCAAGGATAGGAAGCCCTGCCGGCAG 695

Db 182 CTTTCCTCCCTGCACACACACGCTGGGATCTCAAGGATAGGAAGCCCTGCCGGCAG 241

Qy 696 GATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGACAGCAAGAAACACTGT 755

Db 242 GATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGACAGCAAGAAACACTGT 301

Qy 756 ATGTTTTGGAAGCTGCGGAGGACCTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTCTGG 815

Db 302 ATGTTTTGGAAGCTGCGGAGGACCTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTCTGG 361

Qy 816 CCATCAATGCCACAAACATGCTCTGTCTTGTCCCGGAGCCCGCCGCTCTCCCGG 875

Db 362 CCATCAATGCCACAAACATGCTCTGTCTTGTCCCGGAGCCCGCCGCTCTCCCGG 421

Qy 876 CTGCCCACTTGGGAGACGGGTCTTCTGACCTCATCCTCATCCGGAATGCTCCAGGTTCA 935
Db |||||
Qy 422 CTGCCCACTTGGGAGACGGGTCTTCTGACCTCATCCTCATCCGGAATGCTCCAGGTTCA 481
Db |||||
Qy 936 ATTTTCTGAGATTCTCATCAGGCACACCAACGAGCAGGACCAG 979
Db |||||
Qy 482 ATTTTCTGAGATTCTCATCANGCACACCAACGAGCAGGACCAG 525
Db |||||

Search completed: September 6, 2005, 02:19:41
Job time : 3590.6 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 6, 2005, 02:35:14 ; Search time 7073.33 Seconds
(without alignments)
3678.672 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 MGATCAAPLQSLVWVKQQR.....QLVRLFARGIENPKPDHS 537

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10631958/runat.02092005.165815.4533/app_query.fasta_1.1941
-B=GenEmbl -QFMT=fastap -SURFIX=rge -MINMATCH=0.1 -LOOFCLO=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CGN 1.1 11756 @runat.02092005.165815.4533 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 2888 | 100.0 | 1654 | 9 | CR456404 Homo sapi |
| 2 | 2888 | 100.0 | 2042 | 9 | HS457828 Homo sapi |
| 3 | 2888 | 100.0 | 4413 | 6 | AX457006 Sequence |
| 4 | 2888 | 100.0 | 4445 | 9 | AB079066 Homo sapi |

| | | | | | | |
|----|--------|------|--------|----|----------|-----------|
| 5 | 2881 | 99.8 | 4432 | 6 | ARS41900 | Sequence |
| 6 | 2880 | 99.7 | 4463 | 6 | BD102675 | Ceramide |
| 7 | 2688.5 | 93.1 | 1772 | 9 | BC067255 | Homo sapi |
| 8 | 2649.5 | 91.7 | 1840 | 6 | AX224383 | Sequence |
| 9 | 2588 | 89.6 | 4171 | 6 | BD183468 | Novel gen |
| 10 | 2588 | 89.6 | 4171 | 9 | AB051433 | Homo sapi |
| 11 | 2465 | 85.4 | 1459 | 6 | CQ730476 | Sequence |
| 12 | 2436.5 | 84.4 | 2830 | 10 | AB079067 | Mus muscu |
| 13 | 1912.5 | 66.2 | 3661 | 10 | AK129416 | Gallus ga |
| 14 | 1798 | 62.3 | 1450 | 5 | CR386590 | Sequence |
| 15 | 1713 | 59.3 | 2494 | 5 | BC074350 | Xenopus l |
| 16 | 1640.5 | 56.8 | 979 | 6 | AX456998 | Sequence |
| 17 | 1495.5 | 51.8 | 1520 | 5 | BC074110 | Xenopus l |
| 18 | 886.5 | 30.7 | 2333 | 3 | AX112750 | Ciona int |
| 19 | 871 | 30.2 | 550 | 6 | AX457005 | Sequence |
| 20 | 811 | 28.1 | 474 | 6 | AX457001 | Sequence |
| 21 | 801.5 | 27.8 | 2064 | 6 | AR509346 | Sequence |
| 22 | 801.5 | 27.8 | 3178 | 3 | BT014925 | Drosophil |
| 23 | 763 | 26.4 | 1810 | 3 | AY061001 | Drosophil |
| 24 | 710 | 24.6 | 10337 | 6 | CQ608039 | Sequence |
| 25 | 710 | 24.6 | 22353 | 2 | AC017269 | Sequence |
| 26 | 710 | 24.6 | 181009 | 3 | AC007575 | Drosophil |
| 27 | 710 | 24.6 | 295289 | 3 | AE003603 | Drosophil |
| 28 | 675 | 23.4 | 2173 | 6 | CQ608040 | Sequence |
| 29 | 656 | 22.7 | 522 | 6 | AX224385 | Sequence |
| 30 | 648 | 22.4 | 2257 | 8 | AY362552 | Arabidops |
| 31 | 622.5 | 21.6 | 2263 | 8 | AY035126 | Arabidops |
| 32 | 585.5 | 20.3 | 1084 | 6 | AR510195 | Sequence |
| 33 | 578 | 20.0 | 1599 | 9 | AY357073 | Homo sapi |
| 34 | 543 | 18.8 | 3039 | 9 | BC004278 | Homo sapi |
| 35 | 543 | 18.8 | 3074 | 9 | HS459H18 | Homo sapi |
| 36 | 541.5 | 18.8 | 70890 | 9 | HS459H18 | Human DNA |
| 37 | 525 | 18.2 | 329 | 6 | AX457002 | Sequence |
| 38 | 511.5 | 17.7 | 181772 | 2 | AC007689 | Homo sapi |
| 39 | 498 | 17.2 | 26031 | 2 | AC135400 | Rattus no |
| 40 | 450 | 15.6 | 240483 | 2 | AC095955 | Rattus no |
| 41 | 437.5 | 15.1 | 564 | 6 | AX718996 | Sequence |
| 42 | 428 | 14.8 | 1458 | 8 | BT006621 | Arabidops |
| 43 | 428 | 14.8 | 1757 | 8 | AY128394 | Arabidops |
| 44 | 421.5 | 14.6 | 2451 | 9 | AK129976 | Homo sapi |
| 45 | 415 | 14.4 | 1323 | 5 | CR760136 | Xenopus t |

ALIGNMENTS

| | | | | | |
|------------|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|-------------|-------------------------------|
| RESULT 1 | CR456404 | Homo sapiens | bK29F11.1 | full length | open reading frame (ORF) cDNA |
| LOCUS | CR456404 | clone (CDNA clone C22ORF:pGEM.bK29F11.1). | 1654 bp | mRNA | linear PRI 25-MAY-2004 |
| DEFINITION | CR456404 | CR456404.1 | GI:47678338 | | |
| ACCESSION | CR456404 | CR456404.1 | GI:47678338 | | |
| VERSION | CR456404 | CR456404.1 | GI:47678338 | | |
| KEYWORDS | CR456404 | CDNA; chromosome 22; ORF. | | | |
| SOURCE | CR456404 | Homo sapiens (human) | | | |
| ORGANISM | CR456404 | Homo sapiens | | | |
| REFERENCE | CR456404 | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| AUTHORS | CR456404 | 1 (bases 1 to 1654)
Collins, J.E., Wright, C.L., Edwards, C.A., Davis, M.P., Grinham, J.A., Huckle, E.J., Goward, M.E., Aguado, B., Mallia, M., Mokrab, Y., Cole, C.G., Goward, M.E., Aguado, B., Mallia, M., Mokrab, Y., Huckle, E.J., Beare, D.M. and Dunham, I. | | | |
| TITLE | CR456404 | Submitted (24-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: c22@sanger.ac.uk Manuscript | | | |
| JOURNAL | CR456404 | Direct Submission | | | |
| COMMENT | CR456404 | Homo sapiens cDNA sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to isolate cDNA clones representing the full length open reading frame of well annotated protein coding genes on human chromosome 22. For more information see http://www.sanger.ac.uk/HGP/Chr22/.
Location/Qualifiers
1. .1654
/organism="Homo sapiens" | | | |

| | | | | |
|------------------------------------------------------------|-----------|---------------------------------------------------------|----------------------------------------------------------------|------|
| /mol_type="mRNA" | | 201 | SerGluValLeuHisGlyLeuIleGlyValGThrGlnArgSerAlaGlyValAspGlnAsn | 220 |
| /db_xref="taxon:9606" | | 220 | | |
| /chromosome="22" | | 611 | AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGGAGCGCGGGTCGACCAAGAAC | 670 |
| /clone="pGEM.bk29f11.1" | | 221 | HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer | 240 |
| /lab_host="JM109" | | 240 | | |
| 1..1654 | | 671 | CACCCCGGGCTGCTGGTCTCCAGTAGGCTCCGGATTGGAATCATTCGCCGAGGGTCA | 730 |
| /gene="bk29f11.1" | | 730 | | |
| 11..1624 | | 241 | ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis | 260 |
| /gene="bk29f11.1" | | 260 | | |
| /codon_start=1 | | 731 | ACGACACTGCGTGTGTTACTCCACGCTGGGCACACGACGAGAAACCTCGGCCCTGCAT | 790 |
| /protein_id="CAG30290.1" | | 790 | | |
| /db_xref="GI:47678339" | | 261 | IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu | 280 |
| translation="MGATGAAPLQSVLWVKQRCAVSLERPARALLRWRSPPGAGAGA | | 280 | | |
| PGADACSPVSEIIAVBETDVHGKHQSGKQMEKPYAFTVHCVKARRRHWKWAQV | | 280 | | |
| TFWCPERQLCHLWLTQIRELMLEKLTSRPKHLVFINPFGKGQGGKVIYERKVAPLFTL | | 280 | | |
| ASITTDILVTEHANAOKETLYEINIDKVDGIVCVGGDGMSEVLHGLIGRTQPSAGVD | | 280 | | |
| NHTPRAVLVPSLSLIGIIPAGSTDVCYSTVGTSDAETSALHIVGDSLAMDVSSVPH | | 280 | | |
| QNTLLRYSLLRGVFGYDIIKSEKKRWGLARIDFSGLKTLFUSHKCYEGTVTSFLPA | | 280 | | |
| QHTVSPDRKPCRAGFCVCRSQQLLEEQQKALYGLEAAEDVEEWVQVCGKFLAIN | | 280 | | |
| ATNMSCARRSPRGLSAAHLGDGSDLIILIKGSRFNFLRLIHRHTNQDQDFTFV | | 280 | | |
| BYRVKFPQFTSKEMEDSDSLKEGKKRFGHICSSHPSCCTVSNSSWNCNCGBVLHS | | 280 | | |
| PAIEVRVHCQLVRLIFARGIENPKDSSH" | | 280 | | |
| ORIGIN | | 301 | GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu | 320 |
| Alignment Scores: | | 320 | | |
| Pred. No.: | 2,52e-228 | Length: | 1654 | |
| Score: | 2888.00 | Matches: | 537 | |
| Percent Similarity: | 100.00% | Conservative: | 0 | |
| Best Local Similarity: | 100.00% | Mismatches: | 0 | |
| Query Match: | 100.00% | Indels: | 0 | |
| DB: | 9 | Gaps: | 0 | |
| US-10-631-958-10 (1-537) x CR456404 (1-1654) | | 911 | GAGAAGAAACGGTGTGGGTCTTGGCCAGATACACATTTTCAGGTTTAAAGACCTTCTCTC | 970 |
| | | 321 | SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer | 340 |
| | | 340 | | |
| | | 971 | TCCACCACTGCTATGAAGGACAGTGTCTCTCTCTCCCTGCACAAACACACGGTGGATCT | 1030 |
| | | 341 | ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln | 360 |
| | | 360 | | |
| | | 1031 | CCAAGGGATAGGAAGCCCTCGCGGCGAGTGTCTTTGTTGAGGCAAGCAACAGCAGCAG | 1090 |
| | | 361 | LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu | 380 |
| | | 380 | | |
| | | 1091 | CTGAGAGGAGCAGAGAAAGCACTGTATGGTTTGAAGCTGCGAGGAGCTGTGAGGAG | 1150 |
| | | 381 | TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys | 400 |
| | | 400 | | |
| | | 1151 | TGGCAAGTCGTCTGTGGGAAGTTTCTGCCCATCAATGCCACAAACAATGCTCTGTGCTGT | 1210 |
| | | 401 | ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu | 420 |
| | | 420 | | |
| | | 1211 | CGCGAGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTCTTGAGACTC | 1270 |
| | | 421 | IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn | 440 |
| | | 440 | | |
| | | 1271 | ATCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGCAGCACCCNAC | 1330 |
| | | 441 | GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe | 460 |
| | | 460 | | |
| | | 1331 | CAGCAGGACCAAGTTTGACTTCACATTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT | 1390 |
| | | 461 | ThrSerLysHisMetGluAspGluAspSerAspLeuLysGlyLysLysArgPhe | 480 |
| | | 480 | | |
| | | 1391 | ACGTCGAAGCACAATGAGGATGAGCAGCAGCCACTCAAGAGGGGGGGAAGACGCTTT | 1450 |
| | | 481 | GlyHisLysCysSerHisProSerCysCysThrValSerAsnSerSerTrpAsn | 500 |
| | | 500 | | |
| | | 1451 | GGGCACATTTTCAGCAGCCACCCCTCTCTGCTGCTGACCCGCTCTCAACAGCTCTCTGNAAC | 1510 |
| | | 501 | CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal | 520 |
| | | 520 | | |
| | | 1511 | TGCCAGCGGAGGTCCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCGAGCTGGTT | 1570 |
| | | 521 | ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer | 537 |
| | | 537 | | |
| | | 1571 | CGACTCTTTGCAGGAGAAATTGAAGAGAAATCCGAAGCCAGACTCACAGC | 1621 |
| | | 1621 | | |
| RESULT 2 | | | | |
| HSA457828 | | 2042 bp mRNA linear PRI 19-APR-2002 | | |
| LOCUS | | | | |
| DEFINITION | | Homo sapiens mRNA for putative lipid kinase (LK4 gene). | | |

| | | |
|-----------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------|
| ACCESSION | AJ457828 | |
| VERSION | AJ457828.1 | GI:20269072 |
| KEYWORDS | lipid kinase; LK4 gene. | |
| SOURCE | Homo sapiens (human) | |
| ORGANISM | Homo sapiens | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| AUTHORS | 1 | Van Veldhoven, P.P. |
| TITLE | A search for lipid kinases | |
| JOURNAL | Unpublished | |
| REFERENCE | 2 | (bases 1 to 2042) |
| AUTHORS | Van Veldhoven, P.P. | |
| TITLE | Direct Submission | |
| JOURNAL | Submitted (18-APR-2002) Van Veldhoven P.P., Farmakologie, K.U.Leuven, Herestraat, B-3000 Leuven, BELGIUM | |
| FEATURES | Location/Qualifiers | |
| source | 1..2042 | |
| | /organism="Homo sapiens" | |
| | /mol_type="mRNA" | |
| | /db_xref="taxon:9606" | |
| | /chromosome="22" | |
| | /map="22q13" | |
| gene | 1..2042 | |
| | /gene="LK4" | |
| CDS | 71..1684 | |
| | /gene="LK4" | |
| | /codon_start=1 | |
| | /product="putative lipid kinase" | |
| | /protein_id="CAD29884.1" | |
| | /db_xref="GI:20269073" | |
| | /db_xref="GOA:Q8TCT0" | |
| | /db_xref="UniProt/Swiss-Prot:Q8TCT0" | |
| | /translation="MGATGAAPLQSLVMVQRCVASLEPARALLRWRSPGFCAGA PGADCSVPVSEIIIAVEETDVHGKQSGKQMKPYAFTVCHVKRARRHRKWAQV TFWCEPQCHLWLQTLREMLEKLTSPKHLVFINPFGKGQGRKRIYERKVALFTL ASIITDIIVTHANOAKETLYEINIDKYDGLVCVGDGMPSEVLHGLIGRTQRSAGVD QNHPRVLVPSLSRGIIPAGSDCVCSYTGTSDAETSALHIVGDSLAMDVSSVHH NSTLRVSLGLGYFGYDIIKDSKKKWLGLDFSLKTFILSHHCYEGTVSFLPA QHTNSCDRPFRCRAGFCVRSQKQLEEEQKALYGLEAADEVEEVOVQVCGFLAIN ATNMSCARSPRGLSPAHLGDSSDLILRKCSRFNRLRLRHTNQDQDFTFV EYRVKFKQFTSKHMEDESDLKEGGKKRFGHICSSHPSCCCTVSNSSNWCNDCGEVLHS PAIEVRVHCQLVRLPARGIEENPRPDSHS" | |
| ORIGIN | | |
| Alignment Scores: | | |
| Pred. No.: | 3,31e-228 | Length: 2042 |
| Score: | 2888.00 | Matches: 537 |
| Percent Similarity: | 100.00% | Conservative: 0 |
| Best Local Similarity: | 100.00% | Mismatches: 0 |
| Query Match: | 100.00% | Indels: 0 |
| DB: | 9 | Gaps: 0 |
| US-10-631-958-10 (1-537) x HSA457828 (1-2042) | | |
| Qy | 1 | MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrrpVallysGlnGlnArg 20 |
| Db | 71 | ATGGGGGCGACGGGGCGCGAGCCGCTGCAATCCGTGCTGGGTGAAGCAGCAGCGC 130 |
| Qy | 21 | CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrrpArgSerProGlyPro 40 |
| Db | 131 | TGCGCGGTGAGCTGGAGCCCGCGGGCTTCGCTGCGCTGGCGAGCCCGGGGCC 190 |
| Qy | 41 | GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60 |
| Db | 191 | GGAGCGGGCCCGCGCGGAGTCCCTCTCTGTGCGCTGATCTGAGATCATCGCGCTT 250 |
| Qy | 61 | GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrrpGlnLysMetGluLys 80 |
| Db | 251 | GAGGAAACAGACGTTTCAGGGAAACATCAAGGCAGTGGGAAATGCGGAAATGGAAG 310 |
| Qy | 81 | ProTyrAlaPheThrValHisCysVallysArgAlaArgHisArgTrrpAla 100 |

| | | |
|----|------|---------------------------------------------------------------------|
| Db | 311 | CCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGCACCGCTGGAGTGGCGG 370 |
| Qy | 101 | GlnValThrPheTrrpCysProGluGlnLeuCysHisLeuTrrpLeuGlnThrLeuArg 120 |
| Db | 371 | CAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTCTACTTTGTGGTGCAGACCCCTGGCG 430 |
| Qy | 121 | GluMetLeuGluLysLeuThrSerArgProLysHisLeuValPheIleAsnProPhe 140 |
| Db | 431 | GAGATGCTGGGAGAAGCTGACGTCAGACCAAGCATTTTACTGGTATTTATCAACCCGTTT 490 |
| Qy | 141 | GlyGlyLysGlyGlnGlyLysArgIleTrrpGluArgLysValAlaProLeuPheThrLeu 160 |
| Db | 491 | GGAGGAAAGACCAAGCGACGGGATATGAAAGAAAAGTGGCACCACTGTTTCACTTCA 550 |
| Qy | 161 | AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180 |
| Db | 551 | GCCTCCATCACCCTGACATCATGTTACTGAAATGCTTAATCAGGCCAAGGAGACTCTG 610 |
| Qy | 181 | TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200 |
| Db | 611 | TATGAGATTAACTAGACAAATACGACGGCATCGTCTGTGCGGGGAGATGGTATGTTTC 670 |
| Qy | 201 | SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220 |
| Db | 671 | AGCGAGGTGCTGCACGGTCTGATTTGGGAGGACGACGAGGCGCGGGGTGCACCAAGAC 730 |
| Qy | 221 | HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAlaGlySer 240 |
| Db | 731 | CACCCCGGGGTGTGCTGTCCTCCAGTAGCTCCGATTTGGAATCATTTCCCGCAGGGTCA 790 |
| Qy | 241 | ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260 |
| Db | 791 | ACGAGACTCGGTGTGTTACTCCACCGTGGGCACCGACGACGAGAACCTCGCGCGTGCAT 850 |
| Qy | 261 | IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280 |
| Db | 851 | ATCGTTGTGGGGACTCGCTGGCCATGGATGTGCTCCTCAGTCCACCAACACAGCACATC 910 |
| Qy | 281 | LeuArgTyrSerValSerLeuLeuGlyTrrpGlyPheTrrpGlyAspIleLysAspSer 300 |
| Db | 911 | CTTCGCTACTCCGTGTCTCTGCTGGGTACGGCTTCTACGGGGACATCATCAAGGACAGT 970 |
| Qy | 301 | GluLysLysArgTrrpLeuGlyLeuAlaArgTrrpAspPheSerGlyLeuLysThrPheLeu 320 |
| Db | 971 | GAGAAGAAACCGTGTGGGTCTTTCGCCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTC 1030 |
| Qy | 321 | SerHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340 |
| Db | 1031 | TCCACCACTGCTATGAAGGAGCAGTGTCTTCTCTCTGCAACACACACACCGTGGGATCT 1090 |
| Qy | 341 | ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360 |
| Db | 1091 | CAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG 1150 |
| Qy | 361 | LeuGluGluGluGlnLysLysAlaLeuTrrpGlyLeuGluAlaAlaGluAspValGluGlu 380 |
| Db | 1151 | CTGAGAGGAGCAGAGAAACACTGTATGGTTTGGAGCTTGGAGAGCGTGGAGAGAG 1210 |
| Qy | 381 | TrrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400 |
| Db | 1211 | TGGCAAGTCTCTGTGGGAGTTTCTGGCCATCAATGCCACAAACATGTCTCTGTGCTTGT 1270 |
| Qy | 401 | ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420 |
| Db | 1271 | CGCCGAGCCCAAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC 1330 |
| Qy | 421 | IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440 |
| Db | 1331 | ATCCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1390 |
| Qy | 441 | GlnGlnAspGlnPheAspPheThrPheValGluValTrrpArgValLysLysPheGlnPhe 460 |
| Db | 1391 | CAGCAGGACCAAGTTTGACTTCACCTTTTGTGAAAGTTTATTCGCGTCAAGAAATTCAGTTT 1450 |

```
Qy 461 ThrSerLysHisMetGluAspSerAspLeuLysGluGlyGlyLysLysArgPhe 480
|
|
|
Db 1451 ACCTCGAAGACATGGAGGATGAGACACGACCTCAAGAGGGGGGGAAGACGCTTT 1510
|
|
|
Qy 481 GlyHisLysSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
|
|
|
Db 1511 GGGCACATTTGCAGCAGCCACCTCTCTGCTGCTGCACCGCTCTCAACAGCTCTGGAAC 1570
|
|
|
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
|
|
|
Db 1571 TGCAGCGGGAGGTCCTGCACAGCCCTGCCATCAGAGTCCAGCTGCGACGCTGTT 1630
|
|
|
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
|
|
|
Db 1631 CGACTCTTTGCAGGAGGATTAAGAGAAATCCGAAGCCAGACTCACACAG 1681

RESULT 3
AX457006 4413 bp DNA linear PAT 06-JUL-2002
LOCUS
DEFINITION Sequence 9 from Patent WO0228906.
ACCESSION AX457006
VERSION AX457006.1 GI:21715795
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Kossida, S. and Encinas, J.
TITLE Regulation of human sphingosine kinase-like protein
JOURNAL Patent: WO 0228906-A 9 11-APR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source
1..4413
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
76..1689
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38256.1"
/db_xref="GI:21715796"
/translation="MGATGAAPLQSVLVKVKOORCAVSLRPARALLRWSPGPGAGA
PGADACSVPVSEIITAVETDVHKGHKGQWKMEKPYAFVHCVKEARSRHWKQAV
TFWCPERQLCHLNIQTIREMLEKLTSRPKHLVFINPFGKGQKRIYERKVAFLPLT
QNHPRVLVPSSLRIGIIPAGSTDVCYSTVGTSDAETSAHLHIVGDSLAMDVSSVHH
NSTLLRYSVLLGYFGYGDIIKDSEKKRWLGARYDFSLKTLFSLHHCYEGTVSFLPA
QHTVGSPRDRKPRAGCFVCRSQKQLEEEOKKALYGLEAAEDVEEWQVVGKFLAIN
ATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRFNFLIRHTNQQDDPFTV
EYVRVKYQFTSKHMEDESDLKEGGKKRFGHICSSHPSCCTVSNSSNWCNDCGVLHS
PAIEVRVHCOLVRLFARGIENPKDPSH"

ORIGIN
Alignment Scores:
Pred. No.: 8 93e-228 Length: 4413
Score: 2888.00 Matches: 537
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-631-958-10 (1-537) x AX457006 (1-4413)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
|
|
|
Db 76 ATGGGGGCGACGGGGCGCGAGCGCGTGCATCCGTCGTGGTGGTGAAGCAGCAGCGC 135
|
|
|
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuArgTrpTrpArgSerProGlyPro 40
|
|
|
Db 136 TGCACCGGTGAGCCCTGGAGCCCGCGGGGCTCTGCTGGCTGGTGGCGAGCCCGGGGCC 195
|
|
|
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerAspLeu 420
|
|
|
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
|
|
|
Db 196 GGAGCCGGCCGCCCGCGGGATGCTGCTGTGCTGTACTGTAGATCATCGCGTT 255
|
|
|
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
|
|
|
Db 256 GAGGAAACAGACGCTTACCGGGAAACATCAAGGCGAGTGGAAATGGCAGAAATGGAAAG 315
|
|
|
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100
|
|
|
Db 316 CCTTACCGCTTTACAGTTCTCTGTATAAGAGACACGCGCACCCCTGGAAGTGGCG 375
|
|
|
Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
|
|
|
Db 376 CAGGTGACTTTCTGGTGTCCAGAGGACAGCTGTGTCTCTTGGCTGCGACCCCTGGG 435
|
|
|
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuValPheIleAsnProPhe 140
|
|
|
Db 436 GAGATGCTGGAAGAGCTGAGCTCCAGACCAACCAATTTACTGGTATTTATCAACCCGTT 495
|
|
|
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
|
|
|
Db 496 GGAGGAAAGACAAAGCAAGCGGATATATGAAAGAAAGTGGCACCTGTTCCACTTA 555
|
|
|
Qy 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAlaGlnAlaLysGluThrLeu 180
|
|
|
Db 556 GCCTCCATCACTGACATCATCGTTACTGAACATGCTAATCAGGCCAAGGAGACTCTG 615
|
|
|
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
|
|
|
Db 616 TATGAGATTAACTAGACAAATACGACGCATCGTCTGTCTGCGCGGAGATGGTATGTC 675
|
|
|
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
|
|
|
Db 676 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGCGAGGAGCGCGGGGTCGACCAAGAC 735
|
|
|
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
|
|
|
Db 736 CACCCCGGGCTGTGCTGCTCCCAGTAGCTCCGGATTGGAATCATTTCCCGCAGGCTCA 795
|
|
|
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
|
|
|
Db 796 ACGGACTCGCTGTGTTACTCCACCGTGGGCACCGACGACGAGAAACCTCGCGGCTCAT 855
|
|
|
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
|
|
|
Db 856 ATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCACCAACACAGCAGCACTC 915
|
|
|
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
|
|
|
Db 916 CTTGCTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGT 975
|
|
|
Qy 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
|
|
|
Db 976 GAGAAGAAACGGTGTGGGTCTTGGCCAGATACGACTTTTCAGGTTTTAAAGACCTTCCTC 1035
|
|
|
Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
|
|
|
Db 1036 TCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACAGCGGGAATCT 1095
|
|
|
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
|
|
|
Db 1096 CCAAGGGATAGGAAGCCCTCCCGGGCAGGATGCTTTGTTTGCAGGCAAGCAAGCAGCAG 1155
|
|
|
Qy 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
|
|
|
Db 1156 CTGAGAGGAGGAGCAGAAAGACACTGTATGGTTTGAAGCTGCGGAGGACGTTGAGGAG 1215
|
|
|
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
|
|
|
Db 1216 TGGCAAGTCTGTGGGAAGTTTCTGGCCATCATGCCACAAACATGTCTGTGCTGTGT 1275
|
|
|
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerAspLeu 420
|
|
|
```


| | | | |
|------------|------|--------------------------------------------------------------|------|
| Db | 964 | CTTCGCTACTCCGTGTCCTGCTGGCTACGGCTCTACGGGGACATCATCAAGGACAGT | 1023 |
| Qy | 301 | GlulysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu | 320 |
| Db | 1024 | GAGAAGAAACCGTGGTGGGCTTCGCCAGATACACATTTTCAGGTTTAAAGACCTTCCTC | 1083 |
| Qy | 321 | SerHisLysCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer | 340 |
| Db | 1084 | TCCACCACTGCTATGAAGGACAGTGTCTTCTCTCTCCCTGCACACACACGCTGGATCT | 1143 |
| Qy | 341 | ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln | 360 |
| Db | 1144 | CAAAGGATAGCAAGCCCTGCGGCGAGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG | 1203 |
| Qy | 361 | LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu | 380 |
| Db | 1204 | CTGGAGGAGGAGCAAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGGAGCTGGAGAG | 1263 |
| Qy | 381 | TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys | 400 |
| Db | 1264 | TGGCAAGTCGTCTGTGGGAAGTTCTGGCCATCAATGCCACAAACATGTCTGTGCTTGT | 1323 |
| Qy | 401 | ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu | 420 |
| Db | 1324 | CGCCGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGCTCTCTGCACCTC | 1383 |
| Qy | 421 | IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn | 440 |
| Db | 1384 | ATCCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGACACCAAC | 1443 |
| Qy | 441 | GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe | 460 |
| Db | 1444 | CAGCAGGACCAAGTTGACTTCTACTTTTGTGAGTTATCCGGTCAAGAAATTCAGTTT | 1503 |
| Qy | 461 | ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe | 480 |
| Db | 1504 | ACGTGGAAGCACATGAGGATGAGGACAGCGACCTCAAGAGGGGGGGAAGCGCTTT | 1563 |
| Qy | 481 | GlyHisIleCysSerSerHisProSerCysCysThrValSerValSerSerTrpAsn | 500 |
| Db | 1564 | GGGCACATTTGCAGCAGCCACCCCTCCCTGCTGCTGCACCCGCTTCCACAGCTCTGGAAC | 1623 |
| Qy | 501 | CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal | 520 |
| Db | 1624 | TGCGATGGGAGGTCTTCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGCTGGTT | 1683 |
| Qy | 521 | ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer | 537 |
| Db | 1684 | CGACTCTTTGACGAGGAATTAAGAGAATCCGAAGCCAGACTCACACAGC | 1734 |
| RESULT 5 | | | |
| AR541900 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| ACCESSION | | | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| FEATURES | | | |
| source | | | |
| ORIGIN | | | |

| | | | |
|----------------------------------------------|-----------|--------------------------------------------------------------|------|
| Alignment Scores: | | | |
| Pred. No.: | 3.39e-227 | Length: | 4432 |
| Score: | 2881.00 | Matches: | 536 |
| Percent Similarity: | 99.81% | Conservative: | 0 |
| Best Local Similarity: | 99.81% | Mismatches: | 1 |
| Query Match: | 99.76% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |
| US-10-631-958-10 (1-537) x AR541900 (1-4432) | | | |
| Qy | 1 | MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg | 20 |
| Db | 1204 | ATGGGGCGCAGCGGGCGCGGAGCCCGTCAATCCGTGCTGTGGTGAAGACAGCAGCGC | 1263 |
| Qy | 21 | CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro | 40 |
| Db | 1264 | TGCGCCGTGAGCCTGGAGCCCGCGGGGCTCTGCTGCGCTGCTGGCGAGCCCGGGGCC | 1323 |
| Qy | 41 | GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal | 60 |
| Db | 1324 | GGAGCCGGCGCCCGCGCGGATGCCCTGCTCTGTGCTGTATCTGAGATCATCGCGTT | 1383 |
| Qy | 61 | GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys | 80 |
| Db | 1384 | GAGAAACACAGCGTTACGGGAAACATCAAGGCGAGTGGAAATGGCAAAATGGAAGAAG | 1443 |
| Qy | 81 | ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla | 100 |
| Db | 1444 | CTTTACGCTTTACAGTTCTACTGTAAAGAGACACGCGCACCGCTGGAAGTGGCG | 1503 |
| Qy | 101 | GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg | 120 |
| Db | 1504 | CAGGTGACTTCTCGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCCTGCG | 1563 |
| Qy | 121 | GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProphe | 140 |
| Db | 1564 | GAGATGCTGAGAGAGCTGACGTCCAGACCAAGCATTTACTGGTATTTATCAACCCGTTT | 1623 |
| Qy | 141 | GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu | 160 |
| Db | 1624 | GGAGAAAGGACCAAGCAAGCGGATATATGAAGAAAGTGGCACCCACTGTTTCACTTA | 1683 |
| Qy | 161 | AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu | 180 |
| Db | 1684 | GCCTCCATCACTGACATCATCGTTACTGAAATGCTATATCAGGCAAGAGAGACTGTG | 1743 |
| Qy | 181 | TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe | 200 |
| Db | 1744 | TATGAGATTAAATAGACAAATACAGCGGCATCGTCTGTGTCGGCGGAGATGGTATGTTT | 1803 |
| Qy | 201 | SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn | 220 |
| Db | 1804 | AGCAGGTGTGTCACGCTGATTTGGGAGGACGACAGAGGAGCGCGCGGTCGACCAGAAC | 1863 |
| Qy | 221 | HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer | 240 |
| Db | 1864 | CACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGATTCGGATTGGAATCATTTCCCGAGG | 1923 |
| Qy | 241 | ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis | 260 |
| Db | 1924 | ACGAGCTGCGTGTGTTACTCCCGTGGGACACGACGACGAGAAACCTCGCGCTGGAT | 1983 |
| Qy | 261 | IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu | 280 |
| Db | 1984 | ATCGTTGTTGGGACTCGCTGCCCATGGATGTGCTCCTCAGTCCACCAACACAGCACATC | 2043 |
| Qy | 281 | LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer | 300 |
| Db | 2044 | CTTCGCTACTCCGTGTCCCTGCTGCGGTACGGCTTCTACCGGGGACATCATCAAGACAGT | 2103 |
| Qy | 301 | GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu | 320 |
| Db | 2104 | GAGAAGAAACCGTGGTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC | 2163 |

[illegible]


```
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuLysAspSer 300
|||
Db 964 CTTCCGCTACTCCGTGTCCTCGCTGGCTACGGCTTCTACGGGGGACATCATCAAGACAGT 1023
|||
Qy 301 GluLysLysArgTrrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
|||
Db 1024 GAGAAGAAACCGTGTGGGTCTTGCCAGATACGACTTTTTCAGGTCTTAAAGACCTTCCTC 1083
|||
Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
|||
Db 1084 TCCCACCACTCTCTATGAAGGACAGTGTCTTCTCCTCGCACCAACACACGGTGGGATCT 1143
|||
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
|||
Db 1144 CCAAGGATAGAGAGCCCTGCCGGCAGGATGCTTTGTTGTCAGGCCAAGCAACGACGAG 1203
|||
Qy 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
|||
Db 1204 CTGGAGGAGGAGCAGAGAAAGCACTGTATGCTTTTGAAGCTGCGGAGGACGTGGAGGAG 1263
|||
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
|||
Db 1264 TGGCAAGTCGCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTCGTGTGT 1323
|||
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
|||
Db 1324 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCACTTGGAGAGCGGGTCTTCTGCACTC 1383
|||
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
|||
Db 1384 ATCCTCATCCGGAATGCTCCAAGTTCAATTTTCTGAGATTCTCATCAGGCACACCAAC 1443
|||
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
|||
Db 1444 CAGCAGGACCACTTGACTTCACCTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1503
|||
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
|||
Db 1504 ACGTGAGACATCGGAGGATGAGGACAGCAGCCTCAAGAGGGGGGGAAGAGCGCTTT 1563
|||
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
|||
Db 1564 GGGCACATTTGCAGCAGCAGCCCTCTCTGCTGCTGCAACCGTCTCAACAGCTCTGGAAC 1623
|||
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
|||
Db 1624 TGGCATGGGGAGGTCTTCACAGCCCTGCCATCGAGGTGAGAGTCCACTGCCAGCTGGTT 1683
|||
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
|||
Db 1684 CGATCTTTGCAGGGAATTTGAGAGAAATTCGAGACCCAGACTCACAGC 1734
|||
RESULT 7
BC067255 LOCUS
DEFINITION Homo sapiens ceramide kinase, mRNA (cDNA clone IMAGE:6185601), complete cds.
ACCESSION BC067255
VERSION BC067255.1 GI:45595582
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1772)
Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,Z.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
```

```
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 1772)
Strausberg,R.
Direct Submission
Submitted (12-MAR-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh,
Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
Duane Smalilus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 141 Row: d Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21703365
This clone has the following problem: The cds is short compared to
the longest cds in the locus.
Location/Qualifiers
1..1772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6185601"
/tissue_type="Peripheral Nervous System, dorsal root
ganglion"
/clone_lib="Lupski_dorsal_root_ganglion"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..1772
/gene="CERK"
/notes="synonyms: KIAA1646, hCERK, dA59H18.3,
DKFZp344E0211, FLJ21430, FLJ23239, dA59H18.2, LK4"
/db_xref="LocusID:64781"
579..1598
/gene="CERK"
/product_start=1
/product="CERK protein"
/protein_id="AAH67255.1"
/db_xref="GI:45595583"
/db_xref="LocusID:64781"
/translaton="MFSEVLHGLIGTQRSAGVDONHPRAVLVPSSLRIGITIPAGSTD
CVCYSTVGTSDAETSAUHVVGDSLAMDVSSVHNSITLLRYSVLLGYPFGVDLIKDS
EKRWLGLARYDFSLGKLTJSHHCYEGTGFPLPAQHTVSGPRDRKRCAGCFVCRQSK
QOLEEOKKALYGLAEAEVVEWQVCGKFLAINATNATNSCACRRSPRGLSPAHLGDG
```


SSDLILIRKSRFNFELRLIHTNOODQDFTFVEYVRVKKFQFTSKHWEDESDSLKE
GCKKRFHCHSCSHSPCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIENPK
PUSHS"

ORIGIN

Alignment Scores:

Pred. No.: 7, 68e-212 Length: 1772
Score: 2688.50 Matches: 511
Percent Similarity: 94.98% Conservative: 0
Best Local Similarity: 94.98% Mismatches: 1
Query Match: 93.09% Indels: 27
DB: 9 Gaps: 2

US-10-631-958-10 (1-537) x BC067255 (1-1772)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20
Db 60 ATGGGGCGAGCGGGGGCGGAGCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC 119
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 120 TCGCGCGTGAGCTGGAGCCCGCGCGGCTCTGCTGCGCTGTGGCGGAGCCCGGGGCC 179
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db 180 GGAGCGCGCGCGCGCGGATCGCTCTGTGCTGTATCTGAGATCATCGCGCTT 239
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 240 GAGGAAACAGACGTTTACGGGAAACATCAAGCAGTGGGAAATGCGAGAAATGGAAAG 299
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisLeuTrpLysTrpAla 100
Db 300 CCTTACGCTTTTACAGTTTCTCTGTGTAAGAGAGCAGCAGCGCAGCGCTGGAAAGTGGCG 359
Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db 360 CAGGTGACTTCTGTGTGTCAGAGAGAGAGCTGTGTCTGTGCTGCAGACCCCTGCGG 419
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 420 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGGTATTATCAAC----- 473
Qy 141 GlyGlyLys-GlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLe 160
Db 474 -----AAAGGACAAAGGCAAGCGGATATATGAAAGAAAGTGGCACCACCTGTTCACCTT 527
Qy 160 uAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLe 180
Db 528 AGCCTCCATCACCCTGACATCATCGC----- 554
Qy 180 uTyrGluIleAsnIleAspLysTyrAspGlyLeValCysValGlyGlyAspGlyMetPh 200
Db 555 -----ATGCTGTGTGCGGCGAGATGATGTATGTT 583
Qy 200 eSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAs 220
Db 584 CAGCGAGGTGTCACGCTCTGATTGGGAGGACGAGAGGCGCGCGGGTGCACCAAGAA 643
Qy 220 nHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleLeuProAlaGlySe 240
Db 644 CCACCCCGGCTGTGCTGGTCCCGAGTAGCTCCGGATTGGAATCATTCCTCCGAGGGTGC 703
Qy 240 rThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHi 260
Db 704 AACGAGCTGCGTGTGTTACTCCAGCTGGGACCGACGAGCAACCTCGGCGCTGCA 763
Qy 260 sIleValValGlyAspSerLeuAlaMetAspValSerValHisHisAsnSerThrLe 280
Db 764 TATCTGTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCAGCACCACACAGCAGCACT 823
Qy 280 uLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuLysAspSe 300

Db 824 CTTGCGTACTCCTGCTGCTCCCTGCTGGGCTACGGCTTCTTACGGGGACATCATCAAGACAG 883
Qy 300 rGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLe 320
Db 884 TGAGAAGAAACGGTGGTGGTCTTCCAGATACGACATTTTTCAGGTTTAAAGACCTTCTCT 943
Qy 320 uSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySe 340
Db 944 CTCCACCATGCTATGAAGGGACAGTGTCTCTCCCTGTCACAAACACACGCTGGGATC 1003
Qy 340 rProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnG 360
Db 1004 TCCAAGGATAGGAAGCCCTCCGGGAGGATGCTTTGTTGCGAGCAAGACAGCA 1063
Qy 360 nLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGlu 380
Db 1064 GCTGGAGGAGGAGCAGAGAAAGCACGTATGTTTGGAGCTCGGAGGAGCTGGAGGA 1123
Qy 380 uTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCy 400
Db 1124 GTGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCACAAACATGTCTGTGCTTG 1183
Qy 400 sArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLe 420
Db 1184 TCGCGGAGGAGGAGGAGGAGGCTCTCCCGCTGCCACTTTGGGAGACGGGTCTTCTGACCT 1243
Qy 420 uIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAs 440
Db 1244 CATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAA 1303
Qy 440 nGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPh 460
Db 1304 CCAGCAGACCAAGTTTGTGCTTCTTGTGTAAGTTTATCGCGTCAAGAAATTCAGTT 1363
Qy 460 eThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysLeuArgPh 480
Db 1364 TACGTGGAAGCAGCATGGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1423
Qy 480 eGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAs 500
Db 1424 TGGGCACATTTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1483
Qy 500 nCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVa 520
Db 1484 CTGCGAGCGGAGGAGTCTTCACAGCAGCCTGTCATCGAGTCCAGTCCAGTCCAGTGGT 1543
Qy 520 lArgLeuPheAlaArgGlyLeuGluGluAsnProLysProAspSerHisSer 537
Db 1544 TCGACTCTTTGACAGGAGATTTGAAGAGATCCGAAGCCAGACTCACACAGC 1595

RESULT 8
AX224383
LOCUS AX224383 1840 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 5 from Patent WO0160990.
ACCESSION AX224383
VERSION AX224383.1 GI:15554633
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Rastelli, L.
TITLE Novel sphingosine kinases
JOURNAL Patent: WO 0160990-A 5 23-AUG-2001;
Curagen Corporation (US); GENENTECH, INC. (US)
FEATURES
source location/Qualifiers
1. 1840
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Argument Scores: | | | |
| Pred. No.: | 1.32e-208 | Length: | 1840 |
| Score: | 2649.50 | Matches: | 495 |
| Percent Similarity: | 97.8% | Conservative: | 1 |
| Best Local Similarity: | 97.63% | Mismatches: | 0 |
| Query Match: | 91.74% | Indels: | 11 |
| DB: | 6 | Gaps: | 1 |

US-10-631-958-10 (1-537) X AX224383 (1-1840)

| | | | |
|----|-----|------------------------------------------------------------------|-----|
| Qy | 61 | AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaValGlu | 61 |
| Db | 1 | GC CGGCGCCCGCGCGCGATGCCTCTCTGTGCCTGTATCTAGAGATCATCGCGTTGAG | 60 |
| Qy | 62 | GluThrAspValHisGlyIysHisGlnGlySerGlyIysTyrGlnIysMetGluIysPro | 81 |
| Db | 61 | GAACAAGACGTTTACCGGGAAAATCAAGGACGTGGAAATGGCAGAAAATGGAAAAGCCT | 120 |
| Qy | 82 | TyrAlaPheThrValHisCysValIysArgAlaArgHisArgTyrTyrTyrPheAlaGln | 101 |
| Db | 121 | TACGCTTTTACAGTTTACCTGTGTAAAGAGACGACGCGCACCGCTGGAGTGGCGCAG | 180 |
| Qy | 102 | ValThrPheTyrCysProGluGlnIysCysHisLeuTyrTyrGlnIleThrLeuArgGlu | 121 |
| Db | 181 | GTGACTTTCTGTGTGTCCAGAGAGAGAGCTGTGTCACTTTGTGGCTGCAGACCTTCGCGGAG | 240 |
| Qy | 141 | MetLeuGluIysLeuThrSerArgProIysHisLeuLeuValPheIleAsnProPheGly | 141 |
| Db | 241 | ATGCTGGAGAGCTGACGTCCAGACCAAGACATTTTACTGTGTATTATCAACCCGCTTTGGA | 300 |
| Qy | 142 | GlyIysGlyGlnGlyIysArgIleTyrGluArgIysValAlaProLeuPheThrLeuAla | 161 |
| Db | 301 | GGAAAGGACAGCCAGCGGATATATGAAGAAAAGTGGCACCACTGTTCCACTTAGCC | 360 |
| Qy | 170 | SerIleThrThrAspIleIle-----ValThr | 170 |
| Db | 420 | TCCATCACCACCTGACATCATCGGTAAACAAATTCATATGTAACTATGTAGAAGTAATTACT | 420 |
| Qy | 190 | GluHisAlaAsnGlnAlaIysGluThrIleuTyrGluIleAsnIleAspIysTyrAspGly | 190 |
| Db | 480 | GAACATGCTTAATCAGGCCAAGGAGACTCTGTATAGATTAAACATAGACAAATACGACGGC | 480 |
| Qy | 210 | IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyIleGlyArg | 210 |
| Db | 540 | ATCGTCTGTGTCCGCGAGATGGTATGTTCCAGCAGGTGCTGCACGGTCTGATTGGGAGG | 540 |
| Qy | 230 | ThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSer | 230 |
| Db | 600 | ACGCAGAGGAGCGCGGGTCGACAGAACACACCCCGGGCTGTGCTGCTCCCCAGTAGC | 600 |
| Qy | 250 | LeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGly | 250 |
| Db | 660 | CTCCGAGATTGGAAATCATCCCGAGGGTCAACGGACTCGTGTGTACTCCACCGTGGGC | 660 |
| Qy | 270 | ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAsp | 270 |
| Db | 720 | ACCAGCGACGCGAAACCTCGCGCGTGCATATCTGTTTGGGGACTCGCTGGCCATGGAT | 720 |
| Qy | 290 | ValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyr | 290 |
| Db | 780 | GTGTCTTCAGTCCACCACAACGACACATCTCCTTCGCTACTCCGTGTCCTGCTGGGCTAC | 780 |
| Qy | 310 | GlyPheTyrGlyAspIleIleIysAspSerGluIysIysArgTyrTyrLeuGlyIleuAlaArg | 310 |
| Db | 840 | GGCTTCTACGGGGACATCATCAAGGACACGTGAGAAAGAAACGGTGGTGGGTCTTGGCCAGA | 840 |
| Qy | 330 | TyrAspPheSerGlyLeuIysThrPheLeuSerHisHisCysTyrGluGlyThrValSer | 330 |
| Db | 900 | TACGACTTTTTCAGTTTAAAGACCTTCTCTCCCAACCACTGCTATGAAGGACAGTGTCC | 900 |
| Qy | 350 | PheLeuProAlaGlnHisThrValGlySerProArgAspArgIysProCysArgAlaGly | 350 |
| Db | 960 | TTCTCTCCCTGCAACACACGCTGGGATCTCCAAAGGATAGAAAGCCCTGCGCGGACGA | 960 |

| | | | |
|------------|------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------|----------------------------|
| Qy | 351 | CysPheValCysArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyr | 370 |
| Db | 961 | TGCTTTTGGCAGCAAAAGCAGCAGCTGGAGGAGCAGAGAAGACACCTGAT | 1020 |
| Qy | 371 | GlyLeuGluAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAla | 390 |
| Db | 1021 | GGTTTGAAGCTCGGAGGACGTGGAGAGTGGCAAGTCGTGTGGGAAGTTTCGGCC | 1080 |
| Qy | 391 | IleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAla | 410 |
| Db | 1081 | ATCAATGCACAAACATGCTCTGTGTTGTCGGCGAGCCCCAGGGGCTCTCCCGGCT | 1140 |
| Qy | 411 | AlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsn | 430 |
| Db | 1141 | GCCCACTTGGGAGACGGGTCTTCTGACCTCATCTCATCCGGAATGCTCCAGGTTCAAT | 1200 |
| Qy | 431 | PheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheVal | 450 |
| Db | 1201 | TTTCTGAGATTCTCATCAGGCACACCAACAGCAGCAGGTTGACTTCACTTTGTT | 1260 |
| Qy | 451 | GluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAepSer | 470 |
| Db | 1261 | GAAGTTTATCGCTCAAGAAATTCAGTTTACGTGCAAGCACATGGAGGATGAGGACAGC | 1320 |
| Qy | 471 | AspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCys | 490 |
| Db | 1321 | GACCTCAAGAGAGGGGGAAAGACGCTTTGGGCACATTTGCAAGCACCACCTCTCTGC | 1380 |
| Qy | 491 | CysCysThrValSerAsnSerSerTyrAsnCysAspGlyGluValLeuHisSerProAla | 510 |
| Db | 1381 | TGCTGCACCTCTTCCACACAGCTCTGGAACTGGACGGGAGGTCTTGCACAGCCCTGCC | 1440 |
| Qy | 511 | IleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsn | 530 |
| Db | 1441 | ATCAGGTCAAGGTCCACTGCCAGCTGGTTCGACTCTTTCCACAGGAAATTGAAGAGAT | 1500 |
| Qy | 531 | ProLysProAspSerHisSer | 537 |
| Db | 1501 | CCGAAGCCAGACTCACACAGC | 1521 |
| RESULT | 9 | | |
| LOCUS | BD183468 | 4171 bp | DNA linear PAT 17-JUN-2003 |
| DEFINITION | Novel genes and proteins encoded by the genes. | | |
| ACCESSION | BD183468 | | |
| VERSION | BD183468.1 | GI:31875668 | |
| KEYWORDS | JP 2002345492-A/181. | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| AUTHORS | Ohara,O., Nagase,T. and Nakajima,D. | | |
| TITLE | Novel genes and proteins encoded by the genes | | |
| JOURNAL | Patent: JP 2002345492-A 181 03-DEC-2002; | | |
| COMMENT | KAZUSA DNA RESEARCH INSTITUTE | | |
| | OS | Homo sapiens (human) | |
| | PN | JP 2002345492-A/181 | |
| | PD | 03-DEC-2002 | |
| | PF | 26-FEB-2002 | JP 2002049009 |
| | PI | OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA | |
| | PC | C12N15/09,C07K14/47//A61K31/711,A61K38/00,A61K48/00,A61P25/00 | |
| | PC | A61P25/14, | |
| | PC | A61P25/18,A61P35/00,C12N15/00,A61K37/02 | |
| | CC | Novel genes and proteins encoded by the genes | PH |
| | CC | Novel genes and proteins encoded by the genes | Key |
| FEATURES | Location/Qualifiers | | |
| FT | CDS | (2)..(1444). | |
| | Location/Qualifiers | | |
| | 1..4171 | | |
| | /organism="Homo sapiens" | | |
| | /mol_type="genomic DNA" | | |
| | /db_xref="taxon:9606" | | |

ORIGIN

Alignment Scores:

Pred. No.: 4,466-203 Length: 4171
Score: 2588.00 Matches: 481
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.61% Indels: 0
DB: 6 Gaps: 0

US-10-631-958-10 (1-537) x BD183468 (1-4171)

QY 57 llelleAlaValGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGln 76
DB 2 ATCATCGCGGTTGAGGAAACACAGCTTCACGGGAAACATCAAGGCAGTGGGAG 61
QY 77 LysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArg 96
DB 62 AANATGGAAAGCCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGGCAGCC 121
QY 97 TrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuLysHisLeuTrpLeu 116
DB 122 TGGAAAGTGGGCGCAGTGACTTTCTGTGTCTCCAGAGGAGCAGCTGTGTCTGTGGCTG 181
QY 117 GlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuValPhe 136
DB 182 CAGACCCCTGCGGGAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTTACTGTGTA 241
QY 137 lleAenProPheGlyLysGlyGlnGlyLysArgGlyLeuArgLysValAlaPro 156
DB 242 ATCAACCCGTTTGGAGGAAAGGACAGGCAAGCGGATATATGAAGAAAGTGGCACA 301
QY 157 LeuPheThrLeuAlaSerIleThrAspIlelleValThrGluHisAlaAasnGlnAla 176
DB 302 CTGTTCACTTAGCTCCATCACCCTGACATCATCGTTACTGACATGCTTAATCAGGCC 361
QY 177 LysGluThrLeuTyrGluIleAasnIleAspLysTyrAspGlyIleValCysValGlyGly 196
DB 362 AAGGAGACTGTGTATGAGATTAAACATAGACAAATACGACGGCATCGTGTGTGCGGGA 421
QY 197 AspGlyMetPheSerGluValLeuHisGlyLysLeuIleGlyArgThrGlnArgSerAlaGly 216
DB 422 GATGGTATGTTCCAGCGAGTGCTGCACCGTCTGATTTGGAGAGCACAGAGGCCCGGG 481
QY 217 ValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIle 236
DB 482 GTCCAGCAGAACACCCCGGGCTGTGTGTGCTGCCAGTAGCTCCGGATTGGATCAT 541
QY 237 ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr 256
DB 542 CCGCAGGGGTCAACGGACTGCGGTGTGTACTCCACCGTGGGCACGACGCGCAGAAACC 601
QY 257 SerAlaLeuHisIleValValGlyPheSerLeuAlaMetAspValSerSerValHisHis 276
DB 602 TCGGCGTGCATATCGTTTGGGAGCTCGCTGGCCATGGATGTCTCTCAGTCCACCAC 661
QY 277 AenSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 296
DB 662 AACGACACACTCTTCGCTACTCCGCTGCTCGCTGGCTACGGTCTTACGGGACATC 721
QY 297 lleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 316
DB 722 ATCAAGGACAGTGAGAAGAAACGGTGTGGTCTTTCGACATACGACTTTTCAGGTTTA 781
QY 317 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 336
DB 782 AAGACCTTCTCTCCACCACCTGCTATGAAGGAGCAGTGTCTCTCTCCCTGCACAAAC 841
QY 337 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 356
DB 842 ACGTGGGATCTCCAAGGAGATAGGAGCCCTGCGGGCAGGATGTTGTTTGGAGGCA 901
QY 357 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuAlaAlaGlu 376

DB 902 AGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGACACTGTATGTTTGGAAAGCTGCGGAG 961
QY 377 AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet 396
DB 962 GACGTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATG 1021
QY 397 SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 416
DB 1022 TCCTGTGCTTGTGCGCGAGAGCCCGAGGGGCTCTCCCGGCTGCGCACTTGGGAGACGG 1081
QY 417 SerSerAspLeuIleLeuIleArgLysCysSerArgPheAenPheLeuArgPheLeuIle 436
DB 1082 TCCTTCTGACCTCATCTCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATC 1141
QY 437 ArgHisThrAenGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys 456
DB 1142 AGGCACCAACACAGCAGGACAGTTTGACTTCTACTTTTGTGAAGTTTATCGCGTCAAG 1201
QY 457 LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGly 476
DB 1202 AATTTCCAGTTTACGTCGAAGCACATGAGGATGAGGACAGCAGCTCAAGGAGGGGG 1261
QY 477 LysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAen 496
DB 1262 AAGAAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCTCTGCTGCACCGTCTCCAAC 1321
QY 497 SerSerTrpAenCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHis 516
DB 1322 AGCTCTCGGAACCTCGACGGGAGGTCTGCACAGCCCTGCATCGAGTCCAGTCCAC 1381
QY 517 CysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAenProLysProAspSerHis 536
DB 1382 TGCCAGCTGGTTCGACTCTTTGCACGAGGAATGAAGAGAATCGAAGCCAGACTCACAC 1441
QY 537 Ser 537
DB 1442 AGC 1444
RESULT 10
AB051433
LOCUS AB051433 4171 bp mRNA linear PRI 06-OCT-2001
DEFINITION Homo sapiens mRNA for KIAA1646 protein, partial cds.
ACCESSION AB051433
VERSION AB051433.1 GI:13359166
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Hirotsawa,M., Nagase,T., Murahashi,Y., Kikuno,R. and Ohara,O.
TITLE Identification of novel transcribed sequences on human chromosome 22 by expressed sequence tag mapping
JOURNAL DNA Res. 8 (1), 1-9 (2001)
MEDLINE 21156230
PUBMED 11258795
REFERENCE 2 (bases 1 to 4171)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdmainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)
FEATURES
source Location/Qualifiers
1. .4171
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hk01650"
/note="vector:pBluescriptII SK plus"
1. .4171
gene

| CDS | ORIGIN |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|
| /gene="KIAA1646" | |
| <1..1447 | |
| /gene="KIAA1646" | |
| /note="Start codon is not identified." | |
| /codon_start=2 | |
| /product="KIAA1646 protein" | |
| /protein_id="BAB33316.1" | |
| /db_xref="GI:13359167" | |
| /translation="IIAVEETDVHGKHQSGKQMKPYAFTVCHVCARRRRHWKWA
QVTFWCPPEQLCHLMLQTLREMLEKLTSRPKLLVFINPFGKGQKRIYERKVAPLF
TLASITDIITVTHANQAKETLVEINIDKYDGVICVCGDGMFSEVLHGLIGRTQSGV
VDQNHPRVLVPSLRGIIETIPASTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSVF
HHNSTLLRYSGVLSLGYFGDIIKDSKKRWGLARYDFSGELTKTFLSHHCYEGTVSFL
PAQHTVSPDRPCAGFCVQSKQOEKQKALYGLEAAEDVEEQVCGCKFLA
INATNMSCACRRSPRGISPAHLHGDSDDLIRKCSRNFLEFLIRHTNQOQDFPT
FVSVYRKQFTSKMEDESDSLKEGKKRFGHICSSHPSCCTVSNSSWNCDEGLV
HSPAIEVVRVQLVRLPARGIEENPKPDSHS" | |
| Alignment Scores: | |
| Pred. No.: | 4,46e-203 |
| Score: | 2588.00 |
| Percent Similarity: | 100.00% |
| Best Local Similarity: | 100.00% |
| Query Match: | 89.61% |
| DB: | 9 |
| DB: | Gaps: 0 |
| US-10-631-958-10 (1-537) x AB051433 (1-4171) | |
| Qy | 57 IleIleAlaValGluLThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGln 76 |
| Db | 2 ATCATCCCGTTGAGGAAACAGACGTTTACGGGAAACATCAAGGACGTGGAAATGCGAG 61 |
| Qy | 77 LysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArg 96 |
| Db | 62 AAAATGAAAGCCTTACGCTTTTACAGTTTACGTGTGTAAAGAGAGACGACGCGCACCGC 121 |
| Qy | 97 TrpLysTrpAlaGlnValThrPheTrpCysProGluGluLntLeuCysHisLeuTrpLeu 116 |
| Db | 122 TGGNAGTGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTCTCATTGTGGCTG 181 |
| Qy | 117 GlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPhe 136 |
| Db | 182 CAGACCCCTGGGGAGATGCTGGAGAAGCTGACGTCCAGCCAGCAAAAGCATTTACTGGTATT 241 |
| Qy | 137 IleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaPro 156 |
| Db | 242 ATCAACCCGTTTGAGGAAAGGACAGGCAAGCGGATATATGAAAGAAAGTGGCACCA 301 |
| Qy | 157 LeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAla 176 |
| Db | 302 CTGTTACCTTAGCCTCCATCACCACATGACATCATCTGTTACTGAAACATGCTATCAGGCC 361 |
| Qy | 177 LysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGly 196 |
| Db | 362 AAGGAGACTCTGTATGAGTTAAACATAGACAAATACGACGCGCATCGTCTGTGTCGGCGGA 421 |
| Qy | 197 AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly 216 |
| Db | 422 GATGGTATGTTTACGCGAGGTGCTCGACGGTCTGATTGGGAGGACGCGAGAGGAGCGCCGG 481 |
| Qy | 217 ValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIle 236 |
| Db | 482 GTCGACGAGAACCAACCCCGGCGTGTCTGTGTCCAGTAGCTCCGGATTGGTAATCAT 541 |
| Qy | 237 ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr 256 |
| Db | 542 CCCGCGGGTCAACGGACTGCGTGTGTTACTTCCACCGGTGGSCACGACGACGAGAAACC 601 |
| Qy | 257 SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHis 276 |
| Db | 602 TCGCGCTGCATATCGTTGTTGGGGACTCGCTGGCCATGGATGTGCTCTAGTCACCAAC 661 |

| | | | |
|------------|--------------------------------------------------------------------|----------------------------------------------------------------|------|
| Qy | 277 | AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle | 296 |
| Db | 662 | AAcAGACACATCTCCTCGCTACTCCGTGTCCCTCGCTGGGCTACCGCTTCTACGGGGACATC | 721 |
| Qy | 297 | ILeVAspSerGluLysLysArgTyrLeuGlyLeuAlaAlaArgTyrAspPheSerGlyLeu | 316 |
| Db | 722 | ATCAAGGACAGTGAAGAAACGGTGGTGGGTCTTGGCCAGATACGACTTTTCAGGTTTA | 781 |
| Qy | 317 | LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis | 336 |
| Db | 782 | AAGACCTTCTCTCCACCATCTCTATGAAGGACAGTGTCTCTCTCTCCCTGCACAAACAC | 841 |
| Qy | 337 | ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln | 356 |
| Db | 842 | ACGGTGGGATCTCCAAGGATAGAAACCCCTGCCGGCAGGATGCTTTGTTTCAGAGCAA | 901 |
| Qy | 357 | SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu | 376 |
| Db | 902 | AGCAGCAGCAGCTGGAGGAGGAGCAGAAGAACGACTGTATGGTTTGGAAAGCTGCGGAG | 961 |
| Qy | 377 | AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet | 396 |
| Db | 962 | GACGTGGAGAGTGGCAAGTCTGTCTGGGAAAGTTCTGGCCATCAATGCCACAAACATG | 1021 |
| Qy | 397 | SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly | 416 |
| Db | 1022 | TCCTGTGCTTGTCCCGGAGCCCCAGGGGCTCTCCCGGCTGCCACTTTGGGAGACGGG | 1081 |
| Qy | 417 | SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle | 436 |
| Db | 1082 | TCCTTCTGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCTCATC | 1141 |
| Qy | 437 | ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys | 456 |
| Db | 1142 | AGGCACACCACAGCAGGACGAGTTTGACTTTCATCTTTGTTGAGTTTATCGCTCAAG | 1201 |
| Qy | 457 | LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGly | 476 |
| Db | 1202 | AAATTCCAGTTTACGTCGAAGACATCGGAGGATGAGCAGCAGCCCTCAAGAGGGGGGG | 1261 |
| Qy | 477 | LysIleArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsn | 496 |
| Db | 1262 | AAGAAGCGCTTTGGGCACATTTTCAGCAGCAGCACCCCTCTGCTGCTGACCGCTTCCAC | 1321 |
| Qy | 497 | SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHis | 516 |
| Db | 1322 | AGTCTCGGAACATGCAGCGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCAGAGTCCAC | 1381 |
| Qy | 517 | CysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHis | 536 |
| Db | 1382 | TGCCAGCTGGTTCGACTCTTTTGACAGGAGAAATTGAGAGAAATCCGAAGCCAGACTCAC | 1441 |
| Qy | 537 | Ser 537 | |
| Db | 1442 | AGC 1444 | |
| RESULT 11 | | | |
| CQ730476 | | | |
| LOCUS | CQ730476 1459 bp DNA linear PAT 03-FEB-2004 | | |
| DEFINITION | Sequence 16410 from Patent WO02068579. | | |
| ACCESSION | CQ730476 | | |
| VERSION | CQ730476.1 GI:42304409 | | |
| KEYWORDS | Homo sapiens (human) | | |
| SOURCE | Homo sapiens | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| TITLE | Venter, C.J., Adams M.C., Li P.W. and Myers, E.W. | | |
| JOURNAL | Kits, such as nucleic acid arrays, comprising a majority of | | |
| | humanexons or transcripts, for detecting expression and other uses | | |
| | thereof | | |
| | Patent: WO 02068579-A 16410 06-SEP-2002; | | |

PE Corporation (NY) (US)
 Location/Qualifiers
 1. 1459
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1.59e-193 Length: 1459
 Score: 2465.00 Matches: 470
 Percent Similarity: 87.71% Conservative: 1
 Best Local Similarity: 87.52% Mismatches: 6
 Query Match: 85.35% Indels: 61
 DB: 6 Gaps: 3

US-10-631-958-10 (1-537) x CQ730476 (1-1459)

QY 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrrpValLysGlnArg 20
 DB 27 ATGGGGGCGACGGGGCGCGAGCGCTGCAATCCGTGCTGGGTGAAGCAGCAGCGC 86
 QY 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrrpArgSerProGlyPro 40
 DB 87 TCGGCGCAA----- 94
 QY 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
 DB 95 GCAGCTGGTGT---TTTACAGATGCTGCTGTGCTGATATCGAGATCATCGCGTT 151
 QY 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrrpGlnLysMetGluLys 80
 DB 152 GAGGAACAGACGTTTCAGCGGAAACATCAAGGCAGTGGAAAAATGGCAAAATGGAAAG 211
 QY 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrrpLysTrrpAla 100
 DB 212 CTTACGCTTTTACAGTTTACATGTGTAAAGAGCAGCAGCGCCGCTGGAAGTGGCG 271
 QY 101 GlnValThrPheTrrpCysProGluGlnLeuCysHisLeuTrrpLeuGlnThrLeuArg 120
 DB 272 CAGGTGACTTTCTGTGTCCAGAGGAGCAGCTGTCTCACTTGTGCTGCAGACCTCGGG 331
 QY 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
 DB 332 GAGATGCTGGAGAGCTG----- 349
 QY 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
 DB 349 ----- 349
 QY 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
 DB 350 -----ATTACTGAACATGCTAATCAGGCCCAAGGAGACTCTG 385
 QY 181 TyrGluIleAsnIleAspLysTrrpAspGlyLysValGlyCysValGlyAspGlyMetPhe 200
 DB 386 TATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTGCGGCGAGATGTTATGTTTC 445
 QY 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
 DB 446 ACAGAGGTGCTGACGGTCTGATTGGGAGGACGACAGGAGCGCGGGGTGCACAGAAC 505
 QY 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
 DB 506 CACCCCGGGCTGTCTGTTCCAGTAGCTCCGGATTGGATTCATTTCCCGCAGGGTCA 565
 QY 241 ThrAspCysValCysTrrpSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
 DB 566 ACAGACTGCTGTGTACTCCACCGTGGGCACAGCGACGAGAAACCTCGCGCTGTCAT 625
 QY 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
 DB 626 ATCGTTGTGGGACTCGCTGGCCATGATGTCTCCTCAGTCCACCAACAGCAGCAGCTC 685

QY 281 LeuArgTyrSerValSerLeuLeuGlyTrrpGlyPheTyrGlyAspIleIleLysAspSer 300
 DB 686 CTTTCGTACTCCGTGTCCTGCTGGGTACGGCTTCTACGGGACATCATCAAGCAGCAGT 745
 QY 301 GluLysLysArgTrrpLeuGlyLeuAlaArgTrrpAspPheSerGlyLeuLysThrPheLeu 320
 DB 746 GAGAAGAAACGGTGTGGTCTTGGCCAGATACACACTTTTCAGGTTTAAAGACCTTCCTC 805
 QY 321 SerHisHisCysTrrpGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
 DB 806 TCCACCACTGCTATGATGAAGGACAGTGTCTTCTCTCTGCAACAACACACCGTGGGATCT 865
 QY 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
 DB 866 CCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCAAGCAAGCAGCAG 925
 QY 361 LeuGluGluGluGlnLysLysAlaLeuTrrpGlyLeuGluAlaAlaGluAspValGluGlu 380
 DB 926 CTGAGAGGAGGAGCAGAAAGACACTGTATGGTTTGAAGCTGCGGAGGACGTGGAGGAG 985
 QY 381 TrrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
 DB 986 TGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGTGCTGT 1045
 QY 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
 DB 1046 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1105
 QY 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
 DB 1106 ATCTCTATCCGAAATGCTCCAGGTTCATTTCTGAGATTTCTCATCAGCACACCAAC 1165
 QY 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
 DB 1166 CAGCAGACCACTTGTGACTTTCATTTTGTGAAGTTTATCCGCTCAAGAAATTCAGTTT 1225
 QY 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
 DB 1226 ACGTGGAAGCACATGGAGGATGAGCAGCAGCACCTCAAGGAGGGGGGAAGACGCTTT 1285
 QY 481 GlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrrpAsn 500
 DB 1286 GGGCACATTTGCACAGCCACCCCTCTGCTGCTGCACCCGCTCCACAGCTCCTCGGAAC 1345
 QY 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
 DB 1346 TGCACCGGGGAGTCTTCGACAGCCCTGCCATCGAGGTTCAGTCCACTGCGACCTGGTT 1405
 QY 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
 DB 1406 CGACTCTTGCAGGAAATTGAAGAGAAATCCGAAGCAGACTCACAGC 1456

RESULT 12
 AB079067
 LOCUS
 DEFINITION Mus musculus cerk mRNA for ceramide kinases, complete cds.
 ACCSSION
 AB079067
 VERSION AB079067.1 GI:21624341
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Sugiura, M., Kono, K., Liu, H., Shimizu, T., Minekura, H.,
 Spiegel, S., and Kohama, T.
 Ceramide Kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND
 FUNCTIONAL CHARACTERIZATION
 J. Biol. Chem. 277 (26), 23294-23300 (2002)
 JOURNAL
 PUBMED
 11956206
 2 (bases 1 to 2830)
 Sugiura, M., Kono, K., Shimizu, T., Minekura, H., Spiegel, S. and

Kohama, T.
Direct Submission
Submitted (29-JAN-2002) Masako Sugiura, Sankyo Co., Ltd.,
Pharmacology and Molecular Biology Research Laboratories; 2-58
Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan
(E-mail: msugiura@shina.sankyo.co.jp, Tel: 81-3-3492-3131,
Fax: 81-3-5436-8565)
Location/Qualifiers
1. .2830
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1. .2830
/gene="cerk"
1. .1596
/gene="cerk"
/codon_start=1
/product="ceramide kinases"
/protein_id="BAC01155.1"
/db_xref="GI:21624342"
/translation="MGAMGAAPLHSLVWKRRCVSLPRLRRCAVSLPRLRRWRSPPRGP
PGDARSVLSEIIIAVEKDDCEKHASSGRHWKXENPPATVVRKVRVHRHWKWARV
TFWASDQLCHLMLOTURGLLESLSRPHLLVFINPFGGQGGQKRIYKIVAPLFTL
ASITTEIIIEHANQAKETIYEINTSDYGVCGDGMFSEVLHGVIGRTQDSAGID
PNHPRVLVSTLRIIGIPAGSTDCVYSTVGTNDIAETSAHIIIGDSLADVSSVHY
HNTLRYSVSLGYFYGDILIKDSEKKRMGLVRYDPSGLKTFLSHOYREGTSLFLPA
QHTVRSRDNKPCRAGFCVCRSQKOLEEKKALYLENAEEVEEMOVTCGKFLAIN
ATNMSCAPRSPGSLPAHLGGSSDLILIRKSPENFLRIHTNQEDQDFTFV
EVRVKVPHFTSKHVEDNDSRQEKQKFGKICKDRPSCITCSASRSSNMCNDEGVMS
PAIEVRVHCQLVRLFARGIBEEB"

Alignment Scores:
Pred No.: 8.32e-191 Length: 2830
Score: 2436.50 Matches: 454
Percent Similarity: 90.19% Conservative: 33
Best Local Similarity: 84.07% Mismatches: 48
Query Match: 84.37% Indels: 5
DB: 10 Gaps: 1

US-10-631-958-10 (1-537) x AB079067 (1-2830)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20
Db 1 ATGGGGCAATGGGGCGGCGAGCGCTGCACCTCCGTGCTGGTGAACGCGCAGC 60
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 61 TGTCCGCTCAGCTGGAGCCGCGCGGCTCTGTACGCTGGTGGCGAGCCGAGCCC 120
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluLeuAlaVal 60
Db 121 GGGCCCTCCGCGCGGTCGCGATCGCGCTCGGTACTGGTGGAGATCATCGCTGT 180
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 181 GAGGAAAAAGACGACTCGAGAAAAACACGCTCCAGTGGCGGATGGCATATAAATGGAAAAAT 240
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100
Db 241 CCGTTCGCATTCACAGTCCACCGTGTGAAGCGAGTACGACATCACCGCTGGAAGTGGGCA 300
Qy 101 GlnValThrPheTrpCysProGluGlnLeuLeuGlnCysHisLeuTrpLeuGlnThrLeuArg 120
Db 301 CGGTGTACCTTCTGGAGCGCGAGCAGCAGCTGTGTACCTGTGGCTGTGAGACCTCCGT 360
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheLeuAsnProPhe 140
Db 361 GGGCTGTGGAGAGCCTGACTTCAAGACCGAAGCACTTGTGGTATTTCATCAACCCCTTTC 420
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 421 GGAGGGAAGGTGAGGGGACGCGCATCTATGAAAAAACAAGTGGCGGCTCTGTATTACCTTG 480

161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 481 GCTTCCATCACTACGGAGATCATATTACAGACATGCCAACCAAGCAGGAGACTTTA 540
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 541 TACGAGATCAACACACAGACACTATGATGGCATCGTGTGCTTGGTGGGAGCGCATGTC 600
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 601 AGCAGGTGCTGCTGGGTGATTGGAGAGACGAGACAGCGCTGGTATCGACCCCAAT 660
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 661 CACCCCGAGCGCTGCTGGTGGCCAGTACCTCAGGATCGGCATCATACCAGGAGTCC 720
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 721 ACAGATTGTGTGTTACTCAACAGTAGGCACAAACGACGACGAGACATCGGCTTTGCAT 780
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
Db 781 ATCATTATTGGGACCTACCTGGCAATAGACGTGTCTGTGTGCATACCATTAACACGCTG 840
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
Db 841 CTGGGTACTCGGTTCTCTGCTGGCTAGCGTTCTACGGGACCTTAATCAAGGACAGC 900
Qy 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
Db 901 GAAAAGAAACGGTGGATGGGCTCGTCCGGTATGATTCTCAGGGTTGAAGACTTTTCTC 960
Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db 961 TCTCATCAGTACTATGAAGGACACTGTCTCTCCACAGACACACGCGTGGGATCT 1020
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db 1021 CCACGGGACAATAAACCTGCCGGCTGGGTGCTCTGTGTGAGGAGCAGCAAGCAACAG 1080
Qy 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db 1081 CTGGAAGAAGAGAAAGAACCCCTGTATGGCTCGGAGAACCGCGGAAGTGGGAAGAG 1140
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db 1141 TGCAAGTGCATGTGGAAAGTTCCTGGCCATCAATGCCACCAACATGTCTGTGTGTGT 1200
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db 1201 CCTCGGAGCCCTGGGGGCTGTCCCCATTTGGCCATCTGGGAGATGGGTCTTCTGACCTC 1260
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 1261 ATCCTTATCCGGAAGTGTCCAGGTTCACCTTCAGATTTCTCATCCGCGCACACGAAC 1320
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db 1321 CAGGAGACCATGTTCCACTTCTGTTGAAGTTTATCGAGTCAAGAAATTCACACTTC 1380
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
Db 1381 ACCTCGAAGCACCTGGAGATGAGGACATGACTCGAAGGAACAAGAGAGAGCAAGTTT 1440
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerLysSerSerTrpAsn 500
Db 1441 GGAAGATCTCGAAGGACAGACCCCTCTGTGCACTTGCTCAGCCTCCAGACCTCTCTGGAAC 1500
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 1501 TGCATGGCGAAGTCAATGACACGCCCGGCCATTTGAGGTGAGGTCCTCATCTCCAGCTGGTG 1560

```

Qy 521 ArgLeuPheAlaArgGlyIleGluGlu-----AsnProLysProAspSer 535
Db 1561 CGCCTCTTTGCTCGGGGAATCGAGAGAGTCATAGCAAGAACCCCAAGCCCGAGGAGC 1620

RESULT 13
AKI29416 3661 bp mRNA linear ROD 21-NOV-2003
LOCUS Mus musculus mRNA for mKIAA1646 protein.
DEFINITION AKI29416
ACCESSION AKI29416.1 GI:37360495
VERSION FLI CDNA.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
AUTHORS Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Koseki,H., Hiraoka,S.,
Saga,Y., Nagase,T., Ohara,O. and Koga,H.
TITLE Prediction of the coding sequences of mouse homologues of KIAA
gene: iii. the complete nucleotide sequences of 500 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
DNA Res. 10 (4), 167-180 (2003)
JOURNAL 22977043
MEDLINE 14621295
PUBMED 14621295
REFERENCE 2 (bases 1 to 3661)
AUTHORS Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2003) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kanatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:mousse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing: Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
FEATURES
source location/Qualifiers
1..3661
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mph01246"
/tissue_type="embryonic tail"
/note="vector:modified pBC SK+"
1..3661
/gene="mKIAA1646"
<1..1231
/gene="mKIAA1646"
/note="CDS is predicted by in silico analysis. Start codon
is not identified.
Start codon is not identified."
/codon_start=2
/evidence=not experimental
/product="mKIAA1646 protein"
/protein_id="BAC98226.1"
/db_xref="GI:37360496"
IIITEHANQAKETIYEINTSDYGVICVGGDMFSEVLHGIVGRTOQSAGIDPNHRA
VLVPTLRIGIIPAGSTDVCYSTVTGNTDAETSAHIIIGDSLAIIDVSSVHYHNTLLR
YVSLGCGFYGDLLKSEKKRWGLVYDFSGLKTFLSHQYIEGTLSPFLPAQHTVGS
PRDNKPCRGCFVCRKQKEBEKALYGLENAEVEEMQWTCGKFLAINATMNC
ACPRSPGLSPFAHLDGSSDLILIRKCSRFNLFRLIRNQBDFDFTFVEYVRVK
KFHTSHVEDENDSKQEKKQFKGKICKDRPSCSTCSASRSSMNCDEGMHSPAIEVR
VHCQLVRLPARGIBES"

ORIGIN
Alignment Scores:
Pred. No.: 1,81e-147 Length: 3661
Score: 1912.50 Matches: 356
Percent Similarity: 92.49% Conservative: 26
Best Local Similarity: 86.20% Mismatches: 26
Query Match: 66.22% Indels: 5
DB: 10 Gaps: 1

```

```

US-10-631-958-10 (1-537) x AKI29416 (1-3661)
Qy 128 SerArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLys 147
Db 17 TCAAGACCGAAGCACTTGTGTTATTCATCAACCCCTTCGAGGGGAAAGGTCCGGGCAAG 76
Qy 148 ArgIleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIle 167
Db 77 CGCATCTATGAAAAACAGTCGGCCCTCTGTTTACCTTGGCTTCATCACCACACGAGATC 136
Qy 168 IleValThrGluHisAlaAsnGlnAlaValGluThrLeuTyrGluIleAsnIleAspLys 187
Db 137 ATCAATTACAGAGCATGCCAACCAAGCCAGAGAGACTTTATACGAGATCAACACAGACAGC 196
Qy 188 TyrAspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeu 207
Db 197 TATGATGGCATCGTGTGCTAGGTGGGACGGCATGTTTCAGCGAGGTGCTGCATGGGGTG 256
Qy 208 IleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuVal 227
Db 257 ATTGGGAGGACCGACAGACGCGCTGGTATCGACCCCAATCACCCCGAGCGCTGCTGTG 316
Qy 228 ProSerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTyrSer 247
Db 317 CCCAGTACCCCTCAGGATCGGCATCATCCCGCAGGGTCCACAGATTGTGTGTACTCA 376
Qy 248 ThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeu 267
Db 377 ACAGTGGGCACNAACGACGACGACATCGGCTTTCACATCATATTATGGGGACTCGCTG 436
Qy 268 AlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeu 287
Db 437 GCAATAGACGTGCTCTCTGTGCACTACATAAACGCTGCTCGGTACTCTCGTTTCTCTG 496
Qy 288 LeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTyrLeuGly 307
Db 497 CTGGGCTACGGTTTCTACGGGGACTTAATCAAGGACACGCGAAGAACCGGTGATGGGC 556
Qy 308 LeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGly 327
Db 557 CTGCTCGGTATGATTCTCAGGGTGAAGACCTTCTCTCATCATAGTACTATGAAGGG 616
Qy 328 ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCys 347
Db 617 ACCTGTCTCTCCCGACACACACGCTGGGATCTCCACGGGACAAATAAACCCCTGC 676
Qy 348 ArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuLeuGluGluGlnLysLys 367
Db 677 CGGGCTGGGTGCTTCGTGTCGAGCAGACAGCAAGCAGCTGGGAGAGAGAGAGAGAAA 736
Qy 368 AlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLys 387
Db 737 GCCTGTATGGCTGGAGAACGCCGAGGAAGTGAAGAGTGGCAAGTGACATGTGGGAG 796
Qy 388 PheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeu 407
Db 797 TTCCTGGCCATCAATGCCACCAACATGCTCTGTCTGTCTCTCGAGCCCTCGGGGCGCTG 856
Qy 408 SerProAlaAlaHisLeuGlyAspGlySerSerAspIleLeuIleArgLysCysSer 427
Db 857 TCCCATTTGCCATCTCGGAGATGGGTCTTCTGACCTCATCTCTATCCCGAAGTGTCTCC 916
Qy 428 ArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPhe 447
Db 917 AGGTTCAACTTCTGAGATTCTCATCCGCGCACACGAAACCCAGGAGCCAGTTCGACTTC 976
Qy 448 ThrPheValGluValTyrArgValLysPhePheGlnPheThrSerLysHisMetGluAsp 467
Db 977 ACTTTCGTTGAAGTTTATCGAGTCAAGAAATTCACATTCACGTCGAAGACCGTGAAGAC 1036
Qy 468 GluAspSerAspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHis 487

```



```
Db 296 CAATTCCCTTCAGGTGGTGTACCGTCCAGTGACTGAAATGTTTTCAGTAGGGGAAGCA 355
Qy 64 AspValHisGlyIysHisGlnGlySerGlyIysTrpGlnIysMetGluIysProTyrAla 83
Db 356 GAGATTGATGAGAGTACTACAAACAGTATGAAATGGGCAGCATATTTGTAAAGCCTCGTGCA 415
Qy 84 PheThrValHisCysValIysArgAlaArgHisArgTrpLysTyrAlaGlnValThr 103
Db 416 TTTACAGTACGTTATGTACAGAGAGTGGAGAACCCGCTGGCGGTCAAGGAGGTGACC 475
Qy 104 PheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGluMetLeu 123
Db 476 TTTTGGTGCAGTATGAGCAGCTTTTATCAGTGGTTTCAACAGCATTTTCATGACTACTT 535
Qy 124 GluIysLeuThrSerArgProLysHisLeuLeuValPheIleAenProPheGlyClyLys 143
Db 536 GAGCAGCAAACTCATAGACCAAGAACTTGTAGTTTATATTAATCCTTATGTGGGAAG 595
Qy 144 GlyGlnGlyIysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIle 163
Db 596 AAAGAGGTAACAGATATATAGAAATAAAGTGGCTCCATTTATTCAGTGTCTGGGATT 655
Qy 164 ThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIle 183
Db 656 TGTGCTGATGTCATAGTTTACTAGTATGCAACCATGCTAGAGATCACCTATATATGATGCC 715
Qy 184 AsnIleAspIysTyrAspGlyIleValCysValGlyIysValGlyAspGlyMetPheSerGluVal 203
Db 716 AACCTGGAAAAATATGACGGGGTGTGTGTGTGGTGGGATGGAATGTTTCAGCGAAGTG 775
Qy 204 LeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArg 223
Db 776 CTGCATGGCCTCATGTGCAGATGCAAAAGAAATCTGATGTAGATCAACAATCCAGT 835
Qy 224 AlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCys 243
Db 836 GCACAGTTATCTCGGTGTAATATGAGGATGGCATTATCCCTGCTGGTTCAACAGACTGT 895
Qy 244 ValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValVal 263
Db 896 ATATGCTATGCTACTGTGGCATTAATGACCCAGAGAGCTGTCTGTTTGCACATCATTTCTC 955
Qy 264 GlyAspSerLeuAlaMetAspValSerSerValHisIleAsnSerThrLeuLeuArgTyr 283
Db 956 GGTGACTCCAGCCTCTTGATGTATGCTCTGTACACTACAAGAGACATTTCTTAAGTAT 1015
Qy 284 SerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLys 303
Db 1016 TCGGTTTCGTTATTGGGATATGGTTTTTATGGAGATGTTTAAAGGGCAGTGAAAAAAT 1075
Qy 304 ArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHis 323
Db 1076 CGTTGGTTAGGTCCTGTAGATATAGCTTTTCAGGGTTCAAGACATTTCTGACTCACCAT 1135
Qy 324 CysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAsp 343
Db 1136 TGCTATGAAGGGTCAGTCTCATTCAGCCACCAAGTGGGTTCTGGGATCACCAGAGAC 1195
Qy 344 ArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGluGlu 363
Db 1196 CAACAACTTCACATCTGGGTGTACATTTGCAGGCAAGTAGCAAGCAGTAGACGAG 1255
Qy 364 GluGlnLysLysAlaLeuTyrGlyLeuGluAlaGluAspValGluGluTrpGlnVal 383
Db 1256 CAAGAGCAAAACCCAGGCTGTGGATCGAACAACAGAGAAACAGATGACTGGACAACA 1315
Qy 384 ValCysGlyIysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSer 403
Db 1316 ATCAAGAGAAATTCATGGCAATCAATGCTGTACGATGAGTTGTGGCATGTCCACGAAGC 1375
Qy 404 ProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeu 423
Db 1376 CCCAATGGCCCTTTCACACGACGCTCACCTGGCAGATGGAAGTGCAGATCTCATTTTAGTT 1435
```

```
Qy 424 ArgLysCysSerArgPheAenPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAsp 443
Db 1436 AGAAGTGCCTCAAGACTTGATTTTTTAAAGCACCTTATCAGGCACACACCAATAAAGAC 1495
Qy 444 GlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLys 463
Db 1496 CAGTTTGACTTTCCTCAITTTGTTGAAGTTTATCGAGTAAAGAAATTTTCAGTTTACTCCAAA 1555
Qy 464 HisMetGluAspGluAsp-----SerAspLeuLysGluGlyLysLysArgPhe 480
Db 1556 CACTTTGAAGATGAAGACATGAAGACACAGATATAGGAATTTCTGGCAAGAAAACTTT 1615
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAenSerSerTrpAsn 500
Db 1616 TCTCAGATATGCACAGACCATCCATCATGTGGATGTAAACCATGTGATAGCATTTGGAAT 1675
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 1676 TGTGATGGGGAGACATTTGGACCAGACCGCCATTGAAATGAGGTCCACTGCCAACTAATA 1735
Qy 521 ArgLeuPheAlaArgGlyIleGlu 528
Db 1736 AAATTTATTTGCACGTGGCATTGAG 1759
```

Search completed: September 6, 2005, 16:38:47
Job time : 7137.33 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 6, 2005, 02:19:54 ; Search time 829.053 Seconds
(without alignments)
3834.378 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 MGATGAAPLQSLVWVKQR.....QLVRLFARGIEENKPDPSHS 537

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO spool/US10631958/runat 02092005 165814 4497/app query.fasta_1.1941
-DB=N_Geneseq 16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10631958 @CGN 1.1 1436 @runat 02092005 165814 4497 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 16Dec04: *
1: Geneseq1980s: *
2: Geneseq1990s: *
3: Geneseq2000s: *
4: Geneseq2001as: *
5: Geneseq2001bs: *
6: Geneseq2002as: *
7: Geneseq2002bs: *
8: Geneseq2003as: *
9: Geneseq2003bs: *
10: Geneseq2003cs: *
11: Geneseq2003ds: *
12: Geneseq2004as: *
13: Geneseq2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 2888 | 100.0 | 1740 | 8 ADA05679 | Ada05679 Human NOV |
| 2 | 2888 | 100.0 | 4413 | 6 ABL40828 | Ab140828 Human sph |
| 3 | 2888 | 100.0 | 4429 | 12 ADJ96598 | Adj96598 Human lip |
| 4 | 2888 | 100.0 | 4445 | 13 ADP55247 | Adp55247 Human PRO |
| 5 | 2881 | 99.8 | 4432 | 8 ABX70921 | Abx70921 Novel hum |

| | | | | | | |
|----|--------|------|-------|----|----------|---------------------|
| 6 | 2880 | 99.7 | 4463 | 6 | ABA96945 | Aba96945 Human cer |
| 7 | 2843 | 98.4 | 1740 | 12 | ADN62844 | Adn62844 Human NOV |
| 8 | 2649.5 | 91.7 | 1840 | 5 | AAD14426 | Aad14426 Human sph |
| 9 | 2587 | 89.6 | 4231 | 3 | AAA50510 | Aaa50510 Human sph |
| 10 | 2491 | 86.3 | 4702 | 13 | ADS10370 | Adsl0370 Human the |
| 11 | 2349.5 | 81.4 | 2241 | 5 | AAS77728 | Aas77728 DNA encod |
| 12 | 2210 | 76.5 | 3975 | 8 | ACC70838 | Acc70838 Human sph |
| 13 | 1714.5 | 59.4 | 2186 | 5 | AAS77730 | Aas77730 DNA encod |
| 14 | 1640.5 | 56.8 | 979 | 6 | ABL40822 | Ab140822 Human sph |
| 15 | 1386.5 | 48.0 | 1570 | 5 | AAS77731 | Aas77731 DNA encod |
| 16 | 1055 | 36.5 | 753 | 3 | AAC76031 | Aac76031 Human ORF |
| 17 | 1032 | 35.7 | 817 | 6 | ABQ99499 | Abq99499 Human cod |
| 18 | 1032 | 35.7 | 817 | 13 | ADS11669 | Adsl11669 Human the |
| 19 | 986 | 34.1 | 547 | 3 | AAC76592 | Aac76592 Human ORF |
| 20 | 871 | 30.2 | 550 | 6 | ABL40827 | Ab140827 Human sph |
| 21 | 811 | 28.1 | 474 | 6 | ABL40823 | Ab140823 Human sph |
| 22 | 710 | 24.6 | 10337 | 4 | ABL25704 | Ab125704 Drosophil |
| 23 | 679 | 23.5 | 426 | 5 | AAS77727 | Aas77727 DNA encod |
| 24 | 675 | 23.4 | 2173 | 4 | ABL25705 | Ab125705 Drosophil |
| 25 | 656 | 22.7 | 522 | 5 | AAD14427 | Adsl14427 Partial r |
| 26 | 626 | 21.7 | 411 | 5 | AAS77729 | Aas77729 DNA encod |
| 27 | 584 | 20.2 | 1774 | 12 | ADQ88890 | Adq88890 Novel hum |
| 28 | 541.5 | 18.8 | 15185 | 4 | AAK65589 | Aak65589 Human imm |
| 29 | 525 | 18.2 | 329 | 6 | ABL40824 | Ab140824 Human sph |
| 30 | 511.5 | 17.7 | 15181 | 4 | AAK65588 | Aak65588 Human imm |
| 31 | 437.5 | 15.1 | 564 | 8 | ABT23453 | Abt23453 Immune-re |
| 32 | 343.5 | 11.9 | 382 | 5 | AAD14428 | Adsl14428 Partial m |
| 33 | 321 | 11.1 | 2830 | 4 | ABL02327 | Ab102327 Drosophil |
| 34 | 319 | 11.0 | 2609 | 9 | ACF35860 | Acf35860 D. melano |
| 35 | 310.5 | 10.8 | 2020 | 4 | ABL03297 | Ab103297 Drosophil |
| 36 | 310.5 | 10.8 | 2629 | 9 | ACF35859 | Acf35859 D. melano |
| 37 | 310.5 | 10.8 | 4020 | 4 | ABL03296 | Ab103296 Drosophil |
| 38 | 302 | 10.5 | 167 | 6 | ABL40825 | Ab140825 Human sph |
| 39 | 284.5 | 9.9 | 572 | 13 | ACN57851 | Acn57851 Cotton gy |
| 40 | 281 | 9.7 | 2422 | 5 | AAS85331 | Aas85331 DNA encod |
| 41 | 278 | 9.6 | 1857 | 6 | ABK90199 | Abk90199 cDNA enco |
| 42 | 278 | 9.6 | 2380 | 5 | AAS14817 | Aas14817 Human CDN |
| 43 | 278 | 9.6 | 2380 | 6 | ABL59533 | Ab159533 Human sph |
| 44 | 278 | 9.6 | 2380 | 10 | ADE85298 | Ade85298 Farneyyl |
| 45 | 278 | 9.6 | 2380 | 10 | ADE38416 | Ade38416 Human pro |

ALIGNMENTS

| | |
|----------|-----------------------------------------------------------------------|
| RESULT 1 | |
| ADA05679 | |
| ID | ADA05679 standard; cDNA; 1740 BP. |
| XX | |
| AC | ADA05679; |
| XX | |
| DT | 06-NOV-2003 (first entry) |
| XX | |
| XX | Human NOV9a encoding cDNA SEQ ID NO:39. |
| KW | human; NOV9; antidiabetic; anorectic; antibacterial; virucide; |
| KW | immunomodulator; cytostatic; nootropic; neuroprotective; |
| KW | antiparkinsonian; antilipemic; gene therapy; human disease; |
| KW | metabolic disorder; diabetes; obesity; infection; cachexia; cancer; |
| KW | neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; |
| KW | immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FH | Key |
| FT | CDS |
| FT | 76..1689 |
| FT | /*tag= a |
| FT | /product= "NOV9a" |
| XX | |
| PN | WO2003029424-A2. |
| XX | |
| PD | 10-APR-2003. |
| XX | |
| PF | 02-OCT-2002; 2002WO-US031373. |


```
Qy 81 ProTyrAlaPheThrValHisCysValIysArgAlaArgArgHisArgTTrpLysTTrpAla 100
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 316 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGACACGACGCGCACCGCTGGAAGTGGCG 375
Qy 101 GlnValThrPheTTrpCysProGluGluGlnLeuCysHisLeuTTrpLeuGlnThrLeuArg 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 376 CAGGTGACTTTCTCGTGTCCAGAGAGCAGCTGTGTCACTTGTGGCTGCAGACCCCTGCG 435
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 436 GAGATGCTGGAGAGAGCTGACGCTCAGACCAAGCATTTACTGTGTATTATCAACCCCGTTT 495
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 496 GGAGGAAAAGACCAAGCAAGCGGATATATGAAGAAAAGTGGCACCACTGTTTACCTTAA 555
Qy 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 556 GCCTCCATCACCACTGACATCATCGTTTACTGAACATGCTAATCAGGCCAAGGACACTCTG 615
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 616 TATGAGATTAAACATAGACAAATACGACGGCATCGTGTGTGCGCGGAGATGATGTTC 675
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 676 AGCAGGTGCTGCACGGTCTGATTGGGAGGACGACAGGAGCGCGGGGTCCAGCACAAC 735
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 736 CACCCCGGGGTGTGTGTCCTCCAGTAGCCCTCCGGATTGGAATCAITCCCGCAGGGTCA 795
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 796 ACGGACTCGGTGTGTACTCCACCGTGGGACCCAGCGACCGCAGAAACCTCGGCGCTGCAT 855
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 856 ATCGTGTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACACACACAGCACAACCTC 915
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 916 CTTGCTACTCCGTGTCCCTCGCTGGCTTACGGCTTCTACGGGGACATCATCAAGACAGT 975
Qy 301 GluLysLysArgTTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 976 GAGAAGAAACCGTGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC 1035
Qy 321 SerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1036 TCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACCGTGGATCT 1095
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1096 CCAAGGGATAGGAAGCCCTGCGGCGCAGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG 1155
Qy 361 LeuGluGluGlnLysLeuAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1156 CTGAGGAGGAGCAGAAAGAACATGTGTATGTTTGAAGCTGCGGAGGAGCTGGAGGAG 1215
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1216 TGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTGTGCTGT 1275
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1276 CGCCGGAGCCCGCAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC 1335
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1336 ATCTCTATCCCGAAATGCTCCAGGTTCATTTTCTGAGATTCTCATCAGGCACACCAAC 1395
```

```
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1396 CAGCAGGACCAGTTTGTACCTTCACTTTTGTGTAAGTTTATCGCGTCAAGAAATTCACGTTT 1455
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLysLysArgPhe 480
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1456 ACCTCGAAGCACATGGAGGATGAGCAGCAGCGACTCTAAGAGGGGGGGAAGACGCGTTT 1515
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1516 GGGCACATTTGCAGACGCCACCTCTCTGCTGTGACCTCTCCAAACAGCTCTCTGGAAC 1575
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1576 TGGCAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCAAGTCCACTCCAGCTCGGTT 1635
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1636 CGACTCTTTGCAGGAGGATTTGAAGAGAATCCGAAGCCAGACTCACACAGC 1686

RESULT 3
ID ADJ96598 standard; DNA; 4429 BP.
XX
AC ADJ96598;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human lipid kinase KIAA1646 DNA SeqID 55.
XX
KW gene; ds; kinase; human; SNP; single nucleotide polymorphism;
KW tyrosine protein kinase; serine/threonine protein kinase; PTK; STK;
KW gene therapy; cancer; immune-related disease; cardiovascular disease;
KW brain; neuronal associated disease; metabolic; inflammatory disorder;
KW cytosolic; neuroprotective; immunomodulator; antiinflammatory;
KW lipid kinase; KIAA1646.
XX
OS Homo sapiens.
OS 39.
XX
FH Key Location/Qualifiers
FT variation /*tag= a replace(2391,g)
FT /*tag= a /standard_name= "Single nucleotide polymorphism"
FT variation /*tag= b replace(2577,g)
FT /*tag= b /standard_name= "Single nucleotide polymorphism"
FT variation /*tag= c replace(3020,c)
FT /*tag= c /standard_name= "Single nucleotide polymorphism"
FT variation /*tag= d replace(3769,c)
FT /*tag= d /standard_name= "Single nucleotide polymorphism"
FT variation /*tag= e replace(4272,g)
FT /*tag= e /standard_name= "Single nucleotide polymorphism"
XX
PN WO2004006838-A2.
XX
XX
XX 22-JAN-2004.
XX
XX 15-JUL-2003; 2003WO-US021730.
XX
XX 15-JUL-2002; 2002US-0395632P.
XX
XX (SUGS-) SUGEN INC.
XX
XX Whyte D, Manning G, Caenepeel S;
XX
XX WPI; 2004-122753/12.
XX
XX P-PSDB; ADJ96664.
XX
XX New nucleic acid molecule encoding a kinase polypeptide, useful for
```

PT preparing a composition for treating diseases or disorders, e.g., cancer,
 PT or neurological, immunological or inflammatory disorders.

XX Example 1; SEQ ID NO 55; 365pp; English.

PS This invention relates to a novel isolated, enriched or purified nucleic
 XX acid molecule that encodes a kinase polypeptide. Specifically, it relates
 CC to human tyrosine and serine/threonine protein kinases (PK's and STK's),
 CC as well as protein kinase-like enzymes. The present invention describes
 CC screening methods to identify agonists, antagonists and antibodies that
 CC can be used to modulate the activity or function of the mammalian kinase
 CC enzymes. As such, these compositions can be used for gene therapy
 CC purposes to treat diseases or disorders including cancer, immune-related
 CC diseases, cardiovascular disease, brain or neuronal associated disease,
 CC metabolic and inflammatory disorders. Accordingly, they exhibit
 CC cytotostatic, neuroprotective, immunomodulator and antiinflammatory
 CC activities. This polynucleotide sequence is a human kinase DNA sequence
 CC of the invention.

XX
 SQ Sequence 4429 BP; 1020 A; 1115 C; 1213 G; 1081 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,61e-254 Length: 4429
 Score: 2888.00 Matches: 537
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-10-631-958-10 (1-537) x ADJ96598 (1-4429)

QY 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValGlyGlnGlnArg 20
 DB 92 ATGGGGGCGACGGGGCGCGGAGCGCTGCAATCCGTCTGTGGTGAAGCAGCGGC 151
 QY 21 CysAlaValSerLeuGluProAlaAlaGluLeuArgTrpTrpArgSerProGlyPro 40
 DB 152 TGCCTGCGTGGAGCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGCGAGCCCGGGCCC 211
 QY 41 GlyAlaGlyAlaProGlyAlaAlaAspAlaCysSerValProValSerGluIleAlaVal 60
 DB 212 GGAGCGCGCGCCCGCGCGGATGCTGCTCTGTGCTGTATCTGAGATCATCGCGTT 271
 QY 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
 DB 272 GAGGAAACAGACGTTTCAAGGAAACATCAAGCAGTGGGAAATGCGAGAAATGGAAG 331
 QY 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpIleAla 100
 DB 332 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACCGACCGCTGGAAAGTGGCG 391
 QY 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
 DB 392 CAGGTGACTTCTGTGTCCAGAGGAGCAGCTGTGTCTACTGTGCTGCAGACCTCGCG 451
 QY 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
 DB 452 GAGATGCTGGAGAACTGACGCTCCAGACCAAGCATTTTACTGGTATTATCAACCCGTTT 511
 QY 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
 DB 512 GGAGGAAAGGAAAGGAAAGCGGATATATGAAGAAAGTGGCAACCACTGTTCACCTTA 571
 QY 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
 DB 572 GGCTTCATCACCCTGACATCATCTGTTACTGAACTGCTAATCAGGCCAAGGAGACTCTG 631
 QY 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysPglyMetPhe 200
 DB 632 TATGAGATTTAATAGACAAATACGACGCGCATCGTCTGTGTCGGCGGAGATGATGTC 691
 QY 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220

DB 692 AGCAGGTTGCTGCACGGTCTGATTGGGAGGAGCAGCAGGAGCGCGGGGTGCACCAGAAC 751
 QY 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
 DB 752 CACCCCGGGCTGTGCTGGTCCCCCAGTAGCTCCCGATTGGAATCATTTCCCGCAGGGTCA 811
 QY 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
 DB 812 ACGACTGCGTGTGTACTCCACCGTGGCACCAGCAGCAGCAACCTCCGCGCTGCAT 871
 QY 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
 DB 872 ATCGTTGTGGGACTCGCTGGCCATGATGTGCTCCTCAGTCCACCACCAACAGCACATC 931
 QY 281 LeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
 DB 932 CTTCGCTACTCCGTGTCCCTGCTGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT 991
 QY 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
 DB 992 GAGAAGAAACCGTGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC 1051
 QY 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
 DB 1052 TCCACCACTCTATGAAGGACAGTGTCTTCTCCTCGCAACACACACCGTGGGATCT 1111
 QY 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
 DB 1112 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCAAGCAAGCAGCAG 1171
 QY 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
 DB 1172 CTGGAGGAGGAGCAGAAAGAACCTGTATGTTTGGAAAGCTGCGGAGGAGCTGGAGGAG 1231
 QY 381 TrpGlnValValCysGlyLysPheLeuAlaIleAlaThrAsnMetSerCysAlaCys 400
 DB 1232 TGGCAAGTCGTCTGTGGAAAGTTTCTGCCCATCAATGCCAACATGCTCTGTGCTGT 1291
 QY 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
 DB 1292 CGCCGAGCCCAAGGGGCTCTCCCGCGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC 1351
 QY 421 IleIleLysArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
 DB 1352 ATCTCTATCCGAAATGCTCCAGGTTCATTTTCTGAGATTCTCATCAGGACACCAAC 1411
 QY 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
 DB 1412 CAGCAGGACCAAGTTGACTTCACCTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1471
 QY 461 ThrSerLysHisMetGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
 DB 1472 ACGTGAAGCACATGGAGGATGAGACAGCGACCTCAAGGAGGGGGGGAAGAGCGCTTT 1531
 QY 481 GlyHisLysCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
 DB 1532 GGGCACATTTTCAGCAGCAGCCCTCTCTGCTGCTGCACCGCTCCCAACAGCTCTCTGGAAC 1591
 QY 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
 DB 1592 TGCACCGGGAGGTCTGACAGCGCCCTGCCATCAGGTTTCAGAGTCCACTGCCAGCTGGTT 1651
 QY 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
 DB 1652 CGACTCTTGCACGAGGAAATGAAGAAATCCGAAGCCAGACTCACAGC 1702

RESULT 4

ADP55247
 ID ADP55247 standard; cDNA; 4445 BP.
 XX
 AC ADP55247;
 XX
 DT 18-NOV-2004 (first entry)

XX Human PRO cDNA sequence SEQ ID NO:1223.

DE human; PRO; immune related disease; inflammatory immune response;

XX immune response stimulation; antiallergic; antianaemic; antiarthritis;

KW antiasthmatic; antidiabetic; antiinflammatory; antiporiatic;

KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;

KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;

KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;

KW virocid; gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

PN WO2004039956-A2.

XX

PD 13-MAY-2004.

XX

PF 28-OCT-2003; 2003WO-US034381.

XX

PR 29-OCT-2002; 2002US-0422472P.

XX

PA (GETH) GENENTECH INC.

XX

PI Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;

FI Wood WI, Wu TD;

XX

DR WPI; 2004-376182/35.

DR P-PSDB; ADP55248.

XX

XX New PRO polynucleotides and polypeptides, useful in diagnosing

PT and treating an immune related disease, e.g. systemic lupus

PT erythematous, rheumatoid arthritis, diabetes mellitus or asthma and in

PT stimulating an immune response.

XX

PS Claim 2; SEQ ID NO 1223; 3009pp; English.

XX

CC The present invention describes an isolated PRO nucleic acid (1). Also

CC described: (1) a vector comprising (1); (2) a host cell comprising the

CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an

CC isolated PRO polypeptide; (5) a chimeric molecule comprising the

CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an

CC antibody which specifically binds to a polypeptide of (4); (7) a

CC composition of matter comprising a polypeptide of (4), an agonist or

CC antagonist of the polypeptide or an antibody that binds to the

CC polypeptide in combination with a carrier; (8) an article of manufacture

CC comprising a container, a label on the container and a composition of

CC matter of (7); (9) a method of treating an immune related disease in a

CC mammal; (10) a method for determining the presence of a PRO polypeptide

CC in a sample suspected of having the polypeptide; (11) a method of

CC diagnosing an immune related disease or an inflammatory immune response

CC in mammal; (12) a method of identifying a compound that inhibits or

CC mimics the activity of or expression of a gene encoding a PRO polypeptide

CC ; and (13) a method of stimulating the immune response in a mammal. The

CC PRO sequences have antiallergic, antianaemic, antiarthritis,

CC antiasthmatic, antidiabetic, antiinflammatory, antiporiatic,

CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,

CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,

CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and

CC virocid activities, and can be used in gene therapy. The nucleic acid

CC (1) and the encoded polypeptides, compositions, kits and methods are

CC useful in diagnosing and treating an immune related disease and in

CC stimulating an immune response. The present sequence represents a human

CC PRO nucleotide sequence from the present invention.

XX

SQ Sequence 4445 BP; 1018 A; 1134 C; 1216 G; 1077 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 2.63e-254 | Length: | 4445 |
| Scores: | 2888.00 | Matches: | 537 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 13 | Gaps: | 0 |

US-10-631-958-10 (1-537) x ADP55247 (1-4445)

Qy 1 MetClylaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20

Db 124 ATGGGGCGCAGCGGGCGGGAGCCCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC 183

Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40

Db 184 TGGCCGCTGAGCCCTGGAGCCCGCGGGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 243

Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluLeuLeuAlaVal 60

Db 244 GGAGCCGCGCCCGCGCGGGATGCCCTGCTCTGTGCTGTATCTGAGATCATCGCGGTT 303

Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80

Db 304 GAGGAAACAGACGCTTCACGGGAAACATCAAGGGCAGTGGAAATGGCAGAAATGGAAG 363

Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100

Db 364 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGCACCCGCTGGAAGTGGCG 423

Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120

Db 424 CAGGTGACTTTCTGCTGTCAGAGGAGCAGCTGTGTCACTTGTGCTGCAGACCTCGCG 483

Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheLeuAsnProPhe 140

Db 484 GAGATGCTGAGAGAGCTGACGCTCAGACCAAGCATTTACTGTGTTATTTATCAACCCGTT 543

Qy 141 GlyGlyLysGlyGlnGlyLysArgLysValLysValAlaProLeuPheThrLeu 160

Db 544 GGAGGAAAGAGCAGGCAAGCGGATATATGAAGAAAGTGGCACCACCTGTTCACCTTA 603

Qy 161 AlaSerIleThrAspIleLeuValThrGluHisAlaLeuGlnAlaLysGluThrLeu 180

Db 604 GCCTCCATCACCACGACATCATCTGTTACTGAACATGCTAATCAGGCGCAAGGAGACTGTG 663

Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyLysValCysValGlyGlyAspGlyMetPhe 200

Db 664 TATGAGATTAACTAGACAAATACGAGCGCATCTGCTGTGCGCGGAGATGGTATGTTTC 723

Qy 201 SerGluValLeuHisGlyLeuLeuGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220

Db 724 AGCGAGTGTGTCACGCTCTGATTGGGAGGCGCAGAGGAGCGCGCGGGTCGACCAAGAC 783

Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyLysIleProAlaGlySer 240

Db 784 CACCCCGGGCTGTGCTGGTCCCGAGTGGCTCCCGGATTTGGAATCATTTCCCGCAGGGTCA 843

Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260

Db 844 ACGGACTGCGTGTGTACTCCACGCGGGCACCAGCGACGAGAAACCTCGGCGCTGCAT 903

Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280

Db 904 ATCGTTGTTGGGACTCGCTGGCCATGGATGTGCTCTCAGTCCACCAACACAGCAGACTC 963

Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuLysAspSer 300

Db 964 CTTGCTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT 1023

Qy 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320

Db 1024 GAGAGAAACGGTGGTGGGTGTTGCCAGATACACATTTTTCAGGTTTAAAGACCTTCTC 1083

Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340

Db 1084 TCCACCACTGCTATGAGGGACAGTGTCTCTCTCCCTGCACAAACACACGTTGGATCT 1143

Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360

Db 1144 CCAAGGATAGGAGCCCTGCGGCGAGGATGCTTTGTTTCAGGCAAGCAAGCAGAG 1203
Qy 361 LeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 380
Db 1204 CTGGAGGAGGAGCAGAGAAAGCAGCTGTATGCTTTTGGAGCTGCGGAGCGTGGAGG 1263
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db 1264 TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAAATGCTCTGTCTTGT 1323
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db 1324 CGCCGGAGCCCGAGGGCTCTCCCGGCTGCGCACTTGGAGAGCGGTCTTCTGACCTC 1383
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 1384 ATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACCAAC 1443
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyArgValLysLysPheGlnPhe 460
Db 1444 CAGCAGGACCACTTTGACTTCACTTTTGTGTAAGTTTATCGCGTCAAGAAATTTCCAGTT 1503
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
Db 1504 AGCTCGAGGCATCGAGATGAGACAGCGACCTCAAGGAGGGGGGAAGAGCGCTTT 1563
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
Db 1564 GGGCACTTTGAGCAGGACCCCTCTGCTGTCACCGTCTCAACAGCTCTCTGGAAC 1623
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 1624 TCGCATGGGAGGTCTGTCACAGCCCTGCTCATCGAGGTGAGAGTCCACTGCCAGCTGGTT 1683
Qy 521 ArgLeuPheAlaArgGlyLeuGluAsnProLysProAspSerHisSer 537
Db 1684 CGACTCTTTCAGAGGAAATTCAGAGAAATTCAGAGAAATTCAGAGAAATTCAGAGC 1734

RESULT 5

ID ABX70921 standard; cDNA; 4432 BP.

AC ABX70921;

DT 05-MAR-2003 (first entry)

DE Novel human cDNA sequence #146.

XX Human; gene; ss; nervous system disorder; peripheral neuropathy;
KW Huntington's disease; amyotrophic lateral sclerosis; haemophilia;
KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
KW insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;
KW ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;
KW fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;
KW coagulation disorder; cancer; tumour; inflammatory disease; septic shock;
KW Crohn's disease; anaphylaxis; proliferation; chemotactic;
KW differentiation; stem cell growth factor; haematopoiesis; chemoketic;
KW haemostatic; antiinflammatory; expressed sequence tag; EST.

XX Homo sapiens.

XX WO200281731-A2.

XX 17-OCT-2002.

XX 29-JAN-2002; 2002WO-US001222.

XX 30-JAN-2001; 2001US-00774528.

XX (HYSE-) HYSEQ INC.

XX (GOOD/) GOODRICH R W.

PI Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;
XX WPI; 2003-058563/05.
XX Novel polypeptide useful for treating neurodegenerative diseases, myeloid
PT or lymphoid cell disorders, bone disorders, mechanical and traumatic
PT disorders, coagulation disorders, and inflammatory diseases.
XX Claim 1; Page; 612pp; English.
XX This invention relates to the cDNA sequences encoding an isolated novel
CC human polypeptide. The protein encoded by the nucleic acid of the
CC invention is useful for treating central and peripheral nervous system
CC diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic
CC lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,
CC Alzheimer's disease); autoimmune disease (e.g. systemic lupus
CC erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)
CC ; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopaenia)
CC ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,
CC osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head
CC trauma); lung or liver fibrosis; reperfusion injury in various tissues;
CC bacterial, viral or fungal infections; allergic conditions such as
CC allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);
CC cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's
CC disease, anaphylaxis). The protein may be used to inhibit the growth,
CC infection or function of infectious agents such as bacteria, fungi,
CC viruses, or to effect bodily characteristics, biorhythms or circadian
CC cycles of rhythms. The protein may also have
CC proliferation/differentiation, stem cell growth factor, haematopoiesis
CC regulation, immune stimulating or suppressing, chemotactic/chemokinetic,
CC haemostatic and thrombolytic, receptor/ligand, and antiinflammatory
CC activities. The cDNA sequences of the invention are useful for expressing
CC recombinant protein for analysis. The present sequence represents a novel
CC human cDNA sequence of the invention, this sequence is an expressed
CC sequence tag (EST) and was identified using subtractive hybridisation
XX SQ Sequence 4432 BP; 970 A; 1122 C; 1282 G; 1058 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-15e-253 Length: 4432
Score: 2881.00 Matches: 536
Percent Similarity: 99.81% Conservativeness: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 99.76% Indels: 0
DB: 8 Gaps: 0

US-10-631-958-10 (1-537) x ABX70921 (1-4432)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
Db 1204 ATGGGGCGAGCGGGCGGCGGAGCGCTGCAATCCGTGCTGGGTGAAGCAGCAGCGC 1263
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuArgTrpTrpArgSerProGlyPro 40
Db 1264 TGCCTCCGTGAGCCCTGGAGCCGCGCGGCTCTGCTGCGCTGCGGAGCCCGGGGCC 1323
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db 1324 GGAGCCGGGCGCCCGGGGGGATGCTCTGCTGCTGCTGATTCAGATCATCGCGTT 1383
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 1384 GAGGAAACAGACGTTTCAGGGAAACATCAAGGCAGTGGAAATGGCAAAATGGAAAG 1443
Qy 81 ProTyAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100
Db 1444 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGACCGCTGGAAGTGGCG 1503
Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db 1504 CAGGTGACTTCTCTGCTGTCAGAGAGGAGCAGCTGTGCTCCTGCTGCTGCTGCTGCTG 1563

```
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 1564 GAGATGCTGGAGAAAGCTGACGTCCAGACCAAGCATTTACTGTGTTATTTATCAACCCGTTT 1623
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 1624 GGAGGAAAGACAAAGCAAGCGGATATATGAAGAAGAAAGTGGCACCACCTGTTTCACTTA 1683
Qy 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaIleGlnAlaLysGluThrLeu 180
Db 1684 GCCTCCATCACCCTGACATCATCGTTACTGAACTGCTAATCAGGCCAAGGAGACTCTG 1743
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 1744 TATGAGATTACATAGACAATACAGCGCATCGTCTGTGTCGCGGAGATGGTATGTTC 1803
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 1804 AGCAGAGTGTGTCACGGTCTGATTGGGAGGACGAGAGGAGCGCGGGGTCCGACCAAGAC 1863
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 1864 CACCCCCGGGTGCTGTGTCCTCCAGTAGCTCCGGATTGGAATCATTTCCCGCAGGGTCA 1923
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 1924 ACGGACTGCGTGTGTTACTCCACCGTGGCACACGAGCGAGAACTCCGGCGCTGCAT 1983
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSerThrLeu 280
Db 1984 ATCGTTGTGGGAGCTCGCTGGCCATGGATGTCTCCTCAGTCCACCAACACAGCACACTC 2043
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
Db 2044 CTTGCGTACTCCGTGCTCCTCGCTGCGGTACCGCTTCTACGGGGACATCATCAAGGACAGT 2103
Qy 301 GluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
Db 2104 CAGAAGAAACGGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC 2163
Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db 2164 TCCACCACTGCTATGGAGGACAGTGTCTTCTCCTCGCACCAACACACCGTGGGATCT 2223
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db 2224 CCAAGGGATAGAAAGCCCTGCGGCGAGATGCTTTGTTTCAGGCAAGCAACAGCAG 2283
Qy 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db 2284 CTGGAGGAGGAGCAGAAAGCACTGTATGTTTGGAAAGTCTCGGAGGACGTGGAGGAG 2343
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db 2344 TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGTGCTGT 2403
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db 2404 CGCGGAGCCCCAGGGGCTCTCCCGGCTGCCACATTTGGGAGACGGGTCTCTGACCTC 2463
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 2464 ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAAC 2523
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db 2524 CAGCAGGACCAAGTTTGAAGTTTCACTTTTGTAAAGTTTATCGCGTCAAGAAATTCACGTTT 2583
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
Db 2584 ACGTCGAAGCACATGGAGATGAGACAGGACCTCAAGGAGGGGGGAAAGACGCTTT 2643
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
```

```
Db 2644 GGGCACATTTGCAGCAGCCACCCCTCTGCTGTGCTGCACCGTCTCCACAGCTCCTGGAAC 2703
Qy 501 CysAspGlyClyValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 2704 TGGAGCGGGAGGTCTTCGCACGCCCTGCCATCGAGTCCAGATGCCACTGCCAGCTGTT 2763
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProLysProAspSerHisSer 537
Db 2764 CGACTCTTGCACGAGGAATTGAAGAGATCCGAGGCCAGACTCACACAGC 2814

RESULT 6
ABA96945
ID ABA96945 standard; cDNA; 4463 BP.
XX
AC ABA96945;
XX
DT 20-MAY-2002 (first entry)
XX
DE Human ceramide kinase hCERK1-encoding cDNA.
XX
KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;
KW neurological disease; inflammation; human immunodeficiency virus;
KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;
KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;
KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;
KW gene; ss.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 124..1737
XX FT /*tag= a
XX FT /product= "Human ceramide kinase hCERK1"
XX
XX WO200196575-A1.
XX
XX 20-DEC-2001.
XX
XX 11-JUN-2001; 2001WO-JP004889.
XX
XX 14-JUN-2000; 2000JP-00178039.
XX
XX (SANY ) SANKYO CO LTD.
XX
XX Sugiura M, Kono K, Kohama T;
XX
XX WPI; 2002-179513/23.
XX DR P-PSDB; AAM49115.
XX
XX
XX Human ceramide kinase gene and the enzyme encoded by it for screening
XX substances as drugs for neurological, inflammatory and other disorders.
XX
XX Claim 5; Page 46-53; 61pp; Japanese.
XX
XX This sequence represents cDNA encoding a human ceramide kinase designated
XX hCERK1. The invention relates to hCERK1, nucleic acids encoding it,
XX expression vectors and host cells containing hCERK1 nucleic acids, the
XX recombinant production of hCERK1 and antibodies specific for hCERK1. The
XX invention also encompasses methods of isolating hCERK1 from samples, the
XX use of hCERK1 in drug screening, and the use of hCERK1 nucleic acid
XX sequences in gene therapy. hCERK1 mediates the ATP-dependent 1-
XX phosphorylation of ceramides and can be used to screen for therapeutic
XX and preventive agents for a wide range of disorders. Such disorders
XX include neurological disease, inflammation, human immunodeficiency virus
XX (HIV) infection, type 2 diabetes, obesity, sepsis, arteriosclerosis and
XX cancer
XX
XX
XX Sequence 4463 BP; 1026 A; 1135 C; 1217 G; 1084 T; 0 U; 1 Other;
XX
XX
XX Alignment Scores: 1.43e-253 Length: 4463
XX Pred. No.: 2880.00 Matches: 535
XX
XX Score:
```

Percent Similarity: 99.81% Conservative: 1
Best Local Similarity: 99.63% Mismatches: 1
Query Match: 99.72% Indels: 0
DB: 6 Gaps: 0
US-10-631-958-10 (1-537) x ABA96945 (1-4463)
QY 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrrpValLysGlnGlnArg 20
DB 124 ATGGGGGCGAGCGGGGCGGAGCCGCTGCAATCCGCTGTGGGTGAAGCAGCAGCGC 183
QY 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrrpTrrpArgSerProGlyPro 40
DB 184 TGGCCGCTGAGCCTGGAGCCCGCGCGGCTCTGCTGCGCTGTGGCGGAGCCCGGGGCC 243
QY 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
DB 244 GGAGCCGCGCGCCCGCGCGGATGCTGCTGTGCTGTATCTGAGATCATCGCGCTT 303
QY 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrrpGlnLysMetGluLys 80
DB 304 GAGGAAACAGACGTTTACGGGAAACATCAAGGCGAGTGGAAAATGGCAGAAAATGGAAAAG 363
QY 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisAtqTrrpLysTrrpAla 100
DB 364 CCTACGCTTTTACAGTTTACCTGTGTAAGAGAGCAGCAGCCGCTGGAAAGTGGCG 423
QY 101 GlnValThrPheTrrpCysProGluGlnLeuCysHisLeuTrrpLeuGlnThrLeuArg 120
DB 424 CAGGTGACTTCTGCTGTCAGAGGAGCAGCTGTGCTACTTGTGCTGACAGCCCTGCGG 483
QY 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
DB 484 GAGATGCTGGAGAGCTCAGCTCCAGACCAAGCAATTTACTGGTATTTATCAACCCGCTTT 543
QY 141 GlyGlyLysGlyGlnGlyLysArgIleTrrpGluArgLysValAlaProLeuPheThrLeu 160
DB 544 GGAGAAAGAGCAAGCAGCGGATATGAAGAAAAGTGGACCACTGTTTACCTTA 603
QY 161 AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
DB 604 GCCTCCATCACCCTGACATCATCTGTTACTGAACATGCTAATCAGGCCAAGGAGACTCTG 663
QY 181 TyrGluIleAsnIleAspLysTrrpAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
DB 664 TATGAGATTAACATAGACAAATACAGCGCATCTGCTGTGTCGGCGGAGATGTTATGTTT 723
QY 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
DB 724 AGCGAGGTGCTGCAGGCTGATTTGGGAGGACGACAGAGGAGCGCGGGTGCACAGAAC 783
QY 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAlaGlySer 240
DB 784 CACCCCGGGCTGTGCTGCTCCAGTAGCTCCGATTTGGAATCATTTCCCGCAGGTCCA 843
QY 241 ThrAspCysValCysTrrpSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
DB 844 ACGGACTCGTGTGTTACTTCCACCGTGGGACACGACGACGAGAAACCTCGCGCTGCAT 903
QY 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
DB 904 ATCGTGTGTGGGACTCTCGCTGCCATGATGTGCTCAGTCCACCAACAGCAGCACTC 963
QY 281 LeuArgTrrpSerValSerLeuLeuGlyTrrpGlyPheTrrpGlyAspIleLeuAspSer 300
DB 964 CTTCGCTACTCGTGTCTCTGCTGGCTACCGCTTCTACGGGGACATCATCAAGGACAGT 1023
QY 301 GluLysLysArgTrrpLeuGlyLeuAlaArgTrrpAspPheSerGlyLeuLysThrPheLeu 320
DB 1024 GAGAAAGAAACGCTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC 1083
QY 321 SerHisHisCysTrrpGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340

DB 1084 TCCCACTGCTATGAAGGACAGTGTCTCTCTCTCTGACACACACCGTGGATCT 1143
QY 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
DB 1144 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAGCAAGCAGCAG 1203
QY 361 LeuGluGluGlnLysLysAlaLeuTrrpGlyLeuGluAlaAlaGluAspValGluGlu 380
DB 1204 CTGGAGAGGAGCAGAAAGCACTGTATGGTTTGAAGCTGCGGAGGAGCTGGAGGAG 1263
QY 381 TrrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
DB 1264 TGGCAAGTCTGCTGTGGAAAGTTTCTGCCCATCATGCCACAAACATGTCTCTGCTTGT 1323
QY 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
DB 1324 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1383
QY 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
DB 1384 ATCTCTCATCCGGAATGCTCCAAGTTCAATTTCTGAGATTTCTCATCAGGCACCAAC 1443
QY 441 GlnGlnAspGlnPheAspPheThrPheValGluValTrrpArgValLysLysPheGlnPhe 460
DB 1444 CAGCAGACCACTTTGACTTCACTTTTGTGAAGTTTATCGCTCAAGAAATTCAGTTT 1503
QY 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
DB 1504 ACGTGGAAGCACATGGAGGATGAGCAGCAGCCTCAAGGAGGGGGGAGGAGCGCTTT 1563
QY 481 GlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrrpAsn 500
DB 1564 GGGCACATTTTGCAGCAGCCACCCCTCTGCTGTGCTGACCCGCTCTCCACAGCTCTGGAAC 1623
QY 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
DB 1624 TGGATGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTGAGATTCAGTCCACTGCCAGCTGGTT 1683
QY 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
DB 1684 CGACTCTTGCAGGAGAAATTGAAGAGAAATCCGAAGCCAGACTCACACAGC 1734
RESULT 7
ADN62844
ID ADN62844 standard; DNA; 1740 BP.
XX AC ADN62844;
XX AC
XX AC
DT 01-JUL-2004 (first entry)
XX DE Human NOV9a DNA.
XX KW ds; gene; human; NOVX; metabolic disorder; diabetes; obesity;
XX KW infectious disease; anorexia; cancer; cancer-associated cachexia;
XX KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX KW immune disorder; haematopoietic disorder; dyslipidaemia;
XX KW metabolic syndrome X; wasting disorder.
OS Homo sapiens.
XX PN US2004038223-A1.
XX PD 26-FEB-2004.
XX PF 01-OCT-2002; 2002US-00262511.
XX PF 02-OCT-2001; 2001US-0326483P.
XX PR 05-OCT-2001; 2001US-0327435P.
XX PR 05-OCT-2001; 2001US-0327449P.
XX PR 09-OCT-2001; 2001US-0327917P.
XX PR 09-OCT-2001; 2001US-0328029P.
XX PR 09-OCT-2001; 2001US-0328044P.
XX PR 09-OCT-2001; 2001US-0328056P.

| | | | | |
|----|------------------------------------------------------------------------|------------------|----------------------------------------------|---------------------------------------------------------------------------|
| PR | 12-OCT-2001; | 2001US-0328849P. | CC | The invention relates to isolated NOVX polypeptides and polynucleotides. |
| PR | 15-OCT-2001; | 2001US-0329414P. | CC | NOVX polypeptides and polynucleotides are used to prevent, diagnose or |
| PR | 17-OCT-2001; | 2001US-0330142P. | CC | treat a medical condition in human related to the aberrant expression and |
| PR | 18-OCT-2001; | 2001US-0330309P. | CC | activity of NOVX polypeptides. For example, NOVX polypeptides and |
| PR | 22-OCT-2001; | 2001US-0341058P. | CC | polynucleotides may be used to treat disorders associated with decreased |
| PR | 24-OCT-2001; | 2001US-0339266P. | CC | expression or activity of NOVX by supplementing the patient our |
| PR | 24-OCT-2001; | 2001US-0343629P. | CC | production or to rectify mutations. Conversely, antisense NA molecules |
| PR | 29-OCT-2001; | 2001US-0349575P. | CC | may be administered to down regulate expression of NOVX polypeptides by |
| PR | 01-NOV-2001; | 2001US-0346357P. | CC | binding with the cells own genes and preventing their expression. NOVX |
| PR | 17-APR-2002; | 2002US-0373260P. | CC | polynucleotides and complementary sequences may also be used as DNA |
| PR | 19-APR-2002; | 2002US-0373815P. | CC | probes in diagnostic assays to detect and quantitate the presence of |
| PR | 19-APR-2002; | 2002US-0373817P. | CC | similar sequences in samples, and so which patients may be in need of |
| PR | 19-APR-2002; | 2002US-0373826P. | CC | restorative therapy. NOVX polypeptides may also be used as antigens in |
| PR | 19-APR-2002; | 2002US-0373884P. | CC | the production of antibodies and in assays to identify modulators |
| PR | 22-APR-2002; | 2002US-0374977P. | CC | (agonists and antagonists) of the expression and activity of NOVX. The |
| PR | 16-MAY-2002; | 2002US-0381037P. | CC | anti-NOVX polypeptide antibodies, agonists and antagonists may also be |
| PR | 16-MAY-2002; | 2002US-0381038P. | CC | used to modulate NOVX polynucleotide expression and activity of NOVX |
| PR | 16-MAY-2002; | 2002US-0381042P. | CC | polypeptides. The anti-NOVX polypeptide antibodies may also be used as |
| PR | 17-MAY-2002; | 2002US-0381642P. | CC | diagnostic agents for detecting the presence of NOVX in samples. NOVX |
| PR | 28-MAY-2002; | 2002US-0383656P. | CC | polypeptides and polynucleotides may be used in this way to prevent, |
| PR | 29-MAY-2002; | 2002US-0383831P. | CC | diagnose and treat: metabolic disorders, diabetes, obesity, infectious |
| PR | 25-JUN-2002; | 2002US-0391335P. | CC | disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative |
| XX | (SMIT/) | SMITHSON G. | CC | disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, |
| PA | (MILL/) | MILLET I. | CC | haematopoietic disorders, and the various dyslipidaemias, metabolic |
| PA | (PEYM/) | PEYMAN J A. | CC | disturbances associated with obesity, the metabolic syndrome X and |
| PA | (KEKU/) | KEKUDA R. | CC | wasting disorders associated with chronic diseases and various cancers. |
| PA | (JULJ/) | JU J. | CC | They may also be used as antibacterial agents. The present sequence |
| PA | (LILL/) | LI L. | CC | represents DNA encoding a human NOVX protein. |
| PA | (GUOX/) | GUO X. | XX | |
| PA | (PATT/) | PATTURAJAN M. | SQ | Sequence 1740 BP; 388 A; 477 C; 530 G; 345 T; 0 U; 0 Other; |
| PA | (SPYT/) | SPYTEK K A. | | |
| PA | (EDIN/) | EDINGER S R. | Alignment Scores: | |
| PA | (ELLE/) | ELLERMAN K. | Pred. No.: | 9.61e-251 |
| PA | (WALI/) | MALYANKAR U M. | Score: | 2843.00 |
| PA | (ORTT/) | ORT T. | Percent Similarity: | 99.63% |
| PA | (GORM/) | GORMAN L. | Best Local Similarity: | 99.44% |
| PA | (ZERH/) | ZERHUSEN B D. | Query Match: | 12 |
| PA | (ANDE/) | ANDERSON D W. | Indels: | 2 |
| PA | (ZHON/) | ZHONG M. | Gaps: | 0 |
| PA | (CAIT/) | CATTERTON E. | | |
| PA | (JIWW/) | JI W. | US-10-631-958-10 (1-537) x ADN62844 (1-1740) | |
| PA | (MILL/) | MILLER C E. | | |
| PA | (RAST/) | RASTELLI L. | Qy | 1 MetGlyAlaThrGlyAlaAaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20 |
| PA | (STON/) | STONE D J. | Db | 76 ATGGGGGGCGACGGGGGGCGGGAGCCGCTGCAATCCGTGCTGTGGTGAAGACGACGCGC 135 |
| PA | (PENA/) | PENA C E A. | Qy | 21 CysAlaValSerLeuGluProAlaAaArgAlaLeuLeuArgTrpTrpAaSerProGlyPro 40 |
| PA | (SHEN/) | SHENOY S G. | Db | 136 TGGCCCGTGAGCCTGGAGCCCGCGGGGCTCTGCTGCTGTGGCGGAGCCCGGGGGCCC 195 |
| PA | (SHIM/) | SHIMKETS R A. | Qy | 41 GlyAlaGlyAlaProGlyAlaAaSpAlaCysSerValProValSerGluIleAlaVal 60 |
| PA | (LEAC/) | LEACH M D. | Db | 196 GGAGCCGGGGCCCCCGGTGCTGATGCTGCTCTGTGCTGTATCTGAGATCATCGCGTT 255 |
| PA | (AGEE/) | AGEE M L. | Qy | 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80 |
| PA | (BERG/) | BERGHS C. | Db | 256 GAGGAAACAGACGTTTCACGGGAAACATCAAGGCAGTGGAAATGGCAGAAATGGAAAG 315 |
| PA | (DIPI/) | DIPIPPO V A. | Qy | 81 ProTrpAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100 |
| PA | (EISE/) | EISEN A. | Db | 316 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACACCGCGCACCGTGGAGTGGCGG 375 |
| PA | (GANG/) | GANGOLLI E A. | Qy | 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120 |
| PA | (RIEG/) | RIEGER D K. | Db | 376 CAGGTGACTTCTCGTGTCCAGAGGAGCAGCTGTCTACTTGTGGTGCAGACCTCGCG 435 |
| PA | (SPAD/) | SPADERNA S K. | Qy | 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140 |
| XX | Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X; | | Db | 436 GAGATGCTGGAGAAGCTGACGTCCAGACCAAGCAATTTTACTGGTATTTTATCAACCCGTTT 495 |
| PI | Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malynkar UM; | | Qy | 141 GlyGlyLysGlyGlnGlySerArgIleTrpGluArgLysValAlaProLeuPheThrLeu 160 |
| PI | Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E; | | Db | 496 GGAGGAAAAAGGCAAGGCAAGCGGATATATGAAAGAAAGTGGCACCCTGTTTCACTTTA 555 |
| PI | Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG; | | | |
| PI | Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA; | | | |
| PI | Eisen A, Gangolli EA, Rieger DK, Spaderna SK; | | | |
| XX | WPI; 2004-213931/20. | | | |
| DR | P-PSDB; ADN62845. | | | |
| XX | | | | |
| PT | Isolated NOVX polypeptides and nucleic acids, useful for preventing, | | | |
| PT | diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease. | | | |
| XX | Claim 20; SEQ ID NO | | | |

| | | | |
|----|------|---------------------------------------------------------------|------|
| QY | 161 | AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaIleValThrLeu | 180 |
| Db | 556 | GCCTCCATCACCACCTGACATATCGTTACTGAACATGCTTAATCAGGCCAAGAGACTCTG | 615 |
| QY | 181 | TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp-GlyMetPh | 200 |
| Db | 616 | TATGAGATTACATAGACAAATACACGGCAT-GTCTGTGTGGGGAGATCGGTATGTT | 674 |
| QY | 200 | eSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAs | 220 |
| Db | 675 | CAGCGAGGTGCTGCACGGTCTGATTGGGAGAGCAGCAGAGGCGCGGGTCGACCAAA | 734 |
| QY | 220 | nHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySe | 240 |
| Db | 735 | CCACCCCGGGGTGCTGGTCCCGTAGCCTCCGGATTGGAAATCATTCGCGCAGGTC | 794 |
| QY | 240 | rThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHi | 260 |
| Db | 795 | AACGGACTGCGTGTGTTACTCCACGTTGGGCACGACGCGGAAACCTCGGCGCTGCA | 854 |
| QY | 260 | sIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLe | 280 |
| Db | 855 | TATCGTTGTTGGGACTCGTGGCCATGSGATGTGCTCAGTCCACACCACAGCACACT | 914 |
| QY | 280 | uLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSe | 300 |
| Db | 915 | CTTCGCTACTCTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGAC | 974 |
| QY | 300 | rGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLe | 320 |
| Db | 975 | TTAGAAGAAACGGTGGTGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACCTTCT | 1034 |
| QY | 320 | uSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySe | 340 |
| Db | 1035 | CTCCACCACTGCTATGAGGAGCAGTGTCTCTCTCCCTGCACAAACACGGTGGATC | 1094 |
| QY | 340 | rProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGl | 360 |
| Db | 1095 | TCCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTGACGGCAAGCAAGCAGCA | 1154 |
| QY | 360 | nLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGl | 380 |
| Db | 1155 | GCTGAGGAGGAGCAGAAAGCACTGTATGTTTGAAGCTCGGAGGACGTGGAGGA | 1214 |
| QY | 380 | uTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCy | 400 |
| Db | 1215 | GTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTCTGTCTG | 1274 |
| QY | 400 | sArgArgSerProArgGlyLeuSerProAlaAlaHisIleuGlyAspGlySerSerAspLe | 420 |
| Db | 1275 | TGCGCGGAGCCCCAGGGGCTCTCCCGGCTGCCCCACTTGGGAGACGGGTCTTCTGACCT | 1334 |
| QY | 420 | uIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAs | 440 |
| Db | 1335 | CATCTCTCATCCGGAATGCTCAGGTTCAATTTTCTGAGATTTCATCAGGCACACCAA | 1394 |
| QY | 440 | nGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPh | 460 |
| Db | 1395 | CCAGCAGGACGATTTGATCTTCACTTGTGTAAGTTTATCCGTCAGAAATTCACGTT | 1454 |
| QY | 460 | eThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPh | 480 |
| Db | 1455 | TACGTCGAAGCACATGGAGGATGAGCAGCAGCCCTCAAGGAGGGGGGAAGACGCTT | 1514 |
| QY | 480 | eGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAs | 500 |
| Db | 1515 | TGGGCACATTTGAGCAGCAGCCCTCTCTGCTGCTGCACCGCTCTCAACAGCTCTCGAA | 1574 |
| QY | 500 | nCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVa | 520 |
| Db | 1575 | CTGCCACGGGAGGCTCTGCACAGCCCTGCCATCCAGGTCAGAGTCCACTGCCAGCTGTT | 1634 |
| QY | 520 | lArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer | 537 |

Db

1635

TCGACTCTTTGCACGAGGATTTGAAGAGATCCCGAAGCCAGACTCACACAGC

1686

RESULT 8

AAD14426

ID

AAD14426

standard; cDNA; 1840 BP.

XX

AC

AAD14426;

XX

DT

XX

DE

Human sphingosine kinase (SphK) cDNA #2.

XX

KW

Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy; antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis; cytosolic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma; leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss.

KW

OS

Homo sapiens.

XX

Key

Location/Qualifiers

FT

CDS

109..1524

FT

/*tag= a

FT

/product= "Human sphingosine kinase (SphK) protein #2"

XX

PN

WO200160990-A2.

XX

PD

23-AUG-2001.

XX

PF

14-FEB-2001; 2001WO-US004789.

XX

PR

14-FEB-2000; 2000US-0182360P.

PR

22-MAR-2000; 2000US-0191261P.

XX

PA

(CURA-) CURAGEN CORP.

PA

(GETH) GENENTECH INC.

XX

PI

Rastelli L;

XX

DR

WPI; 2001-514770/56.

XX

XX

P-PSDB; AAE07884.

PT

An isolated Sphingosine kinase polypeptide useful for treating a SphK-associated disorder especially cancer, restenosis or ischemia in a human.

XX

PS

Claim 8; Page 94-95; 107pp; English.

XX

CC

The present invention relates to sphingosine kinase (SphK) polypeptides and nucleic acids encoding them. SphK is useful for treating a SphK-associated disorder especially cancers such as leukaemia, lymphoma, ovarian, breast, lung, colon, testicular, stomach and skin, atherosclerosis, restenosis or ischaemia and cell proliferative disease or disorder associated with vascular diseases. SphK gene is used in gene therapy and antisense-therapy. Sphingolipids serving as signalling molecules, have recently emerged as regulators of cell growth, differentiation, diverse cell phenotypes and cell death. Activation of SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human endothelial cells. The present sequence is human sphingosine kinase (SphK) cDNA

XX

SQ

Sequence 1840 BP; 459 A; 468 C; 503 G; 410 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:

6,01e-233

Length:

1840

Score:

2649.50

Matches:

495

Percent Similarity:

97.83%

Conservative:

1

Best Local Similarity:

97.63%

Mismatches:

0

Query Match:

91.74%

Indels:

11

DB:

Gaps:

1

US-10-631-958-10 (1-537) x AAD14426 (1-1840)

QY

42

AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaValGlu

61

[illegible][illegible]

CC sphingosine to form sphingosine 1-phosphate. The polynucleotide was
CC isolated from an HeLa cDNA library by PCR amplification. The invention
CC provides polynucleotides (see AAA50508-10) and polypeptides (see AAY96057
CC -59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC.
CC The polynucleotides may be used as hybridization probes, in the
CC construction of PCR primers for chromosome and gene mapping, in the
CC recombinant production of SKA, SKB and SKC, and in the generation of
CC antisense DNA or RNA. They can be used to detect inflammation or disease
CC associated with abnormal levels of SK expression, or to detect
CC differences in gene sequence between normal and carrier or affected
CC individuals. Host cells expressing SK can be used in drug screening.
CC Human SK specific antibodies, inhibitors, ligands or their analogues are
CC useful as bioactive agents to treat inflammation or disease including
CC viral, bacterial or fungal infections, allergic responses, mechanical
CC injury associated with trauma, hereditary diseases, lymphoma or
CC carcinoma, and other conditions with activate the genes of kidney, lung,
CC heart, lymphoid or tissues of the nervous system
XX

SQ Sequence 4231 BP; 1022 A; 1021 C; 1086 G; 1067 T; 0 U; 35 Other;

Alignment Scores:
Pred. No.: 1.05e-226 Length: 4231
Score: 2587.00 Matches: 481
Percent Similarity: 99.79% Conservative: 1
Best Local Similarity: 99.59% Mismatches: 1
Query Match: 89.58% Indels: 0
DB: 3 Gaps: 0

US-10-631-958-10 (1-537) x AAA50510 (1-4231)

QY 55 SerGluIleAlaValGluThrAspValHisGlyLysHisGlnGlySerGlyLys 74
DB 2 TCTGAGATCATCGCCGTTTGAGGAAACACACGCTTACCGGAAACATCAAGGCGAGTGAATA 61

QY 75 TrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArg 94
DB 62 TGGCAGAAATGGAAGAACCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCG 121

QY 95 HisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLysCysHisLeu 114
DB 122 CACCGCTGGAAGTGGCGCAGGTGACTTCTGTGTCTCCAGAGGAGCAGCTGTGTCACTTG 181

QY 115 TrpLeuGlnThrLeuArgGluMetLeuLysLysLeuThrSerArgProLysHisLeuLeu 134
DB 182 TGGCTGACAGACCTCGCGGAGATGCTGGAGAGCTGACGTCAGACCAACATTTACTG 241

QY 135 ValPheIleAsnProPheGlyLysGlyGlnGlyLysArgIleTyrGluArgLysVal 154
DB 242 GTATTATCAACCGTTTGAGGAAAGGACACAGCAGCGGATATATGAAGAAAGTG 301

QY 155 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsn 174
DB 302 GCACCACTGTTTCACTTACCTCCATCACCCTGACATCATCTGTTACTGAACATGCTAAT 361

QY 175 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 194
DB 362 CAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCTGTGTGC 421

QY 195 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyValTyrGlnArgSer 214
DB 422 GCGGAGATGTATGTTCAGCGAGGTGTGTCCAGCGTCTGATTTGGAGACACGAGGAGC 481

QY 215 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 234
DB 482 GCGGGGTGACACAGAACACCCCGGCTGTGTGTGTCCTCCAGTAGCTCCGGATTGA 541

QY 235 IleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAla 254
DB 542 ATCATTCGCGAGGTCAACGACTGCGTGTGTACTCCACCGTGGGCACGAGCGCA 601

QY 255 GluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerVal 274
DB 602 GAAACCTCGGCGTGCATATCTGTTGGGGACTCGCTGGCCATGGATGTCTCTCAGTC 661

QY 275 HisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGly 294
DB 662 CACCACAACAGCACACTCTTCGCTACTCCCTGTCCTCGTGGGTACTCGGCTTCTACGGG 721

QY 295 AspIleIleLysAspSerGluLysValArgTrpLeuGlyLeuAlaAaTyrAspPheSer 314
DB 722 GACATCATCAAGGACAGTGAGAAACGCTGGTGGTCTTTGCCAGATACGACTTTTCA 781

QY 315 GlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAla 334
DB 782 GGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTCTCCCTGCA 841

QY 335 GlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCys 354
DB 842 CAACACACGGTGGGATCTCCAAGGATAGGAAGCCCTGCCCGGCGAGGATGCTTGTGTC 901

QY 355 ArgGlnSerLysGlnGlnLeuGluGluGlnLysValAlaLeuTyrGlyLeuGluAla 374
DB 902 AGGCAAGCAAGCAGCAGCTGGAGGAGCAGAAAGAACGCTGTATGGTTTTGGAGCT 961

QY 375 AlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThr 394
DB 962 GCGGAGGACGTTGGAGGAGTGGCAAGTCTCTGTGGGAAGTTTCTGGCCATCAATGCCACA 1021

QY 395 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 414
DB 1022 AACATGTCTGTGCTTGTGCCCGGAGCCCGAGGCGCTCTCCCGGCTGCCCACTTGGGA 1081

QY 415 AspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPhe 434
DB 1082 GACGGGTCTTCTGACCTCATCTCTCATCGGAAATGCTCCAGGTTCAATTTTCTGAGATT 1141

QY 435 LeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArg 454
DB 1142 CTGATCAGGCACACCAACACGAGGACGACGTTTCACTTTTGTGAAAGTTTATCGC 1201

QY 455 ValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGlu 474
DB 1202 GTCAAGAAATTCAGTTTACGTCGAAGCACATGAGGAGTATGAGGACGACCTCAAGGAG 1261

QY 475 GlyGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrVal 494
DB 1262 GGGGGGGAAGAGCCCTTTGGGCACATTTGACGAGCAGCACCCCTCTGCTGCACCGTC 1321

QY 495 SerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArg 514
DB 1322 TCCAAAGCTCTGGAAGTGGAGGAGGCTCTGCACAGCCCTGCCATCGAGGTGAGA 1381

QY 515 ValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAsp 534
DB 1382 GTCCACTGCCAGCTGGTTCGACTCTTTGCACGAGATTTGGAAGAGAAATCCGAAGCCAGAC 1441

QY 535 SerHisSer 537
DB 1442 TCACACAGC 1450

RESULT 10
ADS10370
ID ADS10370 standard; DNA; 4702 BP.
XX
AC ADS10370;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human therapeutic DNA - SEQ ID 607.
XX
KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
XX
OS Homo sapiens.
XX

PN WO2004080148-A2.
 XX 23-SEP-2004.
 XX 30-SEP-2003; 2003WO-US030720.
 XX 02-OCT-2002; 2002US-0416186P.
 XX (NUVE-) NUVELO INC.
 XX Tang YT, Asundi V, Ren F, Zhang J, Zhang J, Wehrman T, Wang Z, Ma Y;
 PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AU, Weng G, Zhou P;
 XX WPI; 2004-668857/65.
 DR P-PSDB; ADS11054.
 XX
 PT New polynucleotide, useful in preparing a composition for diagnosing or
 PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
 PT aplastic anemia or cancer for promoting wound healing.
 XX
 PS Claim 1; SEQ ID NO 607; 718pp; English.
 XX
 CC The invention relates to a novel isolated polynucleotide and the encoded
 CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
 CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
 CC be useful in preparing a composition for diagnosing or treating
 CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
 CC disorders, such as aplastic anaemia or cancer, as well as for promoting
 CC wound healing. The molecules may also be utilised during gene therapy
 CC procedures. The current sequence is that of a human therapeutic DNA of
 CC the invention. The current sequence is not shown explicitly within the
 CC specification but can be accessed from the WIPO web-site.
 XX
 SQ Sequence 4702 BP; 970 A; 1226 C; 1364 G; 1118 T; 0 U; 24 Other;

Alignment Scores:
 Pred. No.: 7,86e-218 Length: 4702
 Score: 2491.00 Matches: 493
 Percent Similarity: 65.87% Conservative: 1
 Best Local Similarity: 65.73% Mismatches: 1
 Query Match: 86.25% Indels: 256
 DB: 13 Gaps: 3

US-10-631-958-10 (1-537) x ADS10370 (1-4702)
 QY 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
 DB 818 ATGGGGGGCGACGGGGGGCGGGAGCCGCTGCAATCCGCTGTGGGTGAAGCAGCAGCGC 877
 QY 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
 DB 878 TGGCGCGTGAGCTGGAGCCCGCGGGGCTCTGCTGGCTGTGGCGGAGCCCGGGGCC 937
 QY 41 GlyAlaGlyAlaProGlyAla----- 47
 DB 938 GGAGCGGGCGCCCGGGCGC-GGCGCTACCAAGGCCAGCCCTCTTCCAGACTCACCT 996
 QY 47 ----- 47
 DB 997 AGCACTGGTTATTTTGGCAGTATGTCCAAAGACTGGCGAGCGTTTACACCTTCTTTAGC 1056
 QY 47 ----- 47
 DB 1057 TGGTTCTCTGGCACTGCTCCAGCAGGAAACGGCGCTGGTACTGCCAGGGCGCTGCACCC 1116
 QY 47 ----- 47
 DB 1117 ACCCTGTGCCCCATAGCTCAGAGCTCCCGAGGCCAGAGCTGGAGATAAACCCAGCTCC 1176
 QY 47 ----- 47
 DB 1177 CAGCTGTGACTCTGCACGCTGGGCTTCTCCAGGCCAGCGGGCGCTGCTGTGGATCTCTG 1236

QY 47 ----- 47
 DB 1237 GTGTCCCGGTTCACTGCGCTGTTTCTTGTATCTTGTTCAGAGCTATCAGTCATAGCGGA 1296
 QY 47 ----- 47
 DB 1297 CGGAGGCTGATGACAGGTTTCCCGGACCTGACAGCATGAGTGGGACCGCGGGTGTTC 1356
 QY 48 -----AspAlaCysSerValProVal 54
 DB 1357 TGTGTTGAGATGCCAGCTGAAGGACGTGGTGTGTTTACAGATGCTGTCCTGTCCTGTA 1416
 QY 55 SerGluIleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLys 74
 DB 1417 TCTGAGATCATCGCGTTGAGGAAACAGACGTTTACGGGAAACATCAAGCAGTGGAAAA 1476
 QY 75 TrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArg 94
 DB 1477 TGGCAGAAAAATGGAAAAAGCCTTACGCTTTTACAGTTTCACTGTCTAAAGAGAGCAGCAGG 1536
 QY 95 HisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeu 1114
 DB 1537 CACCGCTGGAAAGTGGCGCAGGTGACTTCTGTGTCCAGAGGAGCAGCTGTGTCACTTG 1596
 QY 115 TrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeu 134
 DB 1597 TGGCTGCAGACCCCTGCGGGAGATGCTGGAGAAGCTG----- 1632
 QY 135 ValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysVal 154
 DB 1632 ----- 1632
 QY 155 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsn 174
 DB 1633 -----ATTACTGAACATGCTAAT 1650
 QY 175 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 194
 DB 1651 CAGGCCAAGGAGACTCTGTATGAGATTAACATAGACAAATACGACGCGCATCGTGTGTGC 1710
 QY 195 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSer 214
 DB 1711 GCGCGAGATGTTATGTTTCAGCAGGTGCTGCACGGTCTGATTGGGAGGACGCGAGAGGC 1770
 QY 215 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 234
 DB 1771 GCGGGGTCCACAGAACACACCCCGGGCTGTGCTGCTCCAGTAGCTCCGATTTGA 1830
 QY 235 IleIleProAla----- 238
 DB 1831 ATCATTTCCCGCAGGCTTTTGCAATGACCTGGCAGGGGACGAGGTGTCTGTCTCTTGGC 1890
 QY 238 ----- 238
 DB 1891 CCTGTGTGTCGCCCGAGGGTGGCGCANTGGTGACACTTTTCACTCTCGTCACAGCTCTG 1950
 QY 238 ----- 238
 DB 1951 GGATGTGAGCACCGCAGTCAATCCCATTTTATGGATGAACAGAGGAGACTGGGGAGCAT 2010
 QY 239 -----GlySerThrAspCysValCysTyrSer 247
 DB 2011 GTGGGCCCCGGTGAGAACGCTGTGGCTTGGACGGGTCAACGGAGTCCGCTGTGTACTCC 2070
 QY 248 ThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeu 267
 DB 2071 ACCGTGGGCGCCAGCAGCAGCAAAACCTCGCGCTGCATATCGTTGTGTGGGAGCTCCGCTG 2130
 QY 268 AlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeu 287
 DB 2131 GCCATGGATGTCTCTCAGTCCACCACCAACAGCACACTCTCTCGCTACTCCGTGTCCTG 2190
 QY 288 LeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGly 307


```
Db 2191 CTGGGCTACGGCTTCTACGGGACATCATCAGGACAGTGAGAGAAACGGTGTGGGT 2250
Qy LeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGly 327
Db 2251 CTTCGCAGATACGACTTTTCAGGTTTAAAGACCTTCCTCCACCACCTGCTATGAGGG 2310
Qy 328 ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCys 347
Db 2311 ACAGGTGCTCTCCCTCCCTCCCAACACACAGCGTGGGATCTCAAGGGATAGAGCCCTGC 2370
Qy 348 ArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLysLys 367
Db 2371 CGGGCAGATGCTTTGTTTGCAGGCAAGACAGCAGCTGAGGAGGAGCAGAGAA 2430
Qy 368 AlaLeuTyrGlyLeuGluAlaGluAspValGluGluTrpGlnValValCysGlyLys 387
Db 2431 GCACGTGTATGTTTGAAGCTCGGAGGACGTGGAGGAGTGCGAAGTCTGTGTGGGAAG 2490
Qy 388 PheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeu 407
Db 2491 TTTCTGGCCATCAATGCCACAAACATGCTCTGTGCTGTCGCGAGCCCCAGGGGCTC 2550
Qy 408 SerProAlaHisLeuGlyAspGlySerSerAspLeuIleLeuArgLysCysSer 427
Db 2551 TCCCGGCTGCCACTTGGAGACGGTCTTCTGACCTCATCTCCGAAATGCTCC 2610
Qy 428 ArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPhe 447
Db 2611 AGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCACTTTGACTTC 2670
Qy 448 ThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAsp 467
Db 2671 ACTTTTGTGAAGTTTATTCGGCTCAAGAAATTCACAGTTTACGTGCAAGCACAATGAGGAT 2730
Qy 468 GluAspSerAspLeuLysGluGlyCysLysArgPheGlyHisIleCysSerSerHis 487
Db 2731 GAGGACGACCTCAAGAGGGGGGGAAGCGCTTTGGCCATTTGAGCAGGCCAC 2790
Qy 488 ProSerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHis 507
Db 2791 CCTCTCTGCTGTCACCGTCTCCAAACAGCTCCTGGAACTCGGACGGGGAGTCTTGAC 2850
Qy 508 SerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIle 527
Db 2851 AGCCCTGCATCAGAGTCAGAGTCACATGCCAGCTGTTTCGACTCTTTGCACGAGGAAT 2910
Qy 528 GluGluAsnProLysProAspSerHisSer 537
Db 2911 GAAGAGAAATCCGAAGCCAGACTCACACAGC 2940

RESULT 11
AAS77728
ID AAS77728 standard; cDNA; 2241 BP.
AC AAS77728;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #13532.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
```

```
PR 23-AUG-2000; 2000US-00649167.
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-FSDB; ABG13541.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 13532; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridization probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2241 BP; 453 A; 627 C; 700 G; 461 T; 0 U; 0 Other;

Alignment Scores:
Pred No.: 2,73e-205 Length: 2241
Score: 2349.50 Matches: 471
Percent Similarity: 64.50% Conservative: 5
Best Local Similarity: 81.82% Mismatches: 8
Query Match: 81.35% Indels: 255
DB: Gaps: 5

US-10-631-958-10 (1-537) x AAS77728 (1-2241)
Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
Db 83 ATGGGGCGCAGCGGGCGCGGAGCCGCTGCAATCCGTGCTGGGTGAACGAGCAGCGC 142
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 143 TGCGCCGTGAGCCCTGGAGCCCGCGCGGCTCTGCTGCGCTGCTGCGAGCCCGGGGCC 202
Qy 41 GlyAlaGlyAlaProGlyAla----- 47
Db 203 GGAGCCGGCGCCCGCGCGCGGCGGCGGCGGCTCACCAGGCCCGCCCTCTTCCAGACTCACCT 261
Qy 47 ----- 47
Db 262 AGCACTGTTATTTTGCAGTATGTTCCAAGACTGGCGAGCGGTTTACACCTTCTGTTTAGC 321
Qy 47 ----- 47
Db 322 TGGTTTCTCTGGCAGCTGCTCCAGCAGGGAACGGCGCTGGTACTGCCAGGGCGCTGCACCC 381
Qy 47 ----- 47
Db 382 ACCCTGCTGCCCCNATAGCTCAGACGTCGCCGAGGCCAGAGCTGGAGATAAACCCAGCTCC 441
```

```
Qy 47 ----- 47
Db 442 CAGCTGTGACTCTGCAGCGTGGCCTCTCCAGCCAGCGGGCCTGCTGCTGGATCTCTG 501
Qy 47 ----- 47
Db 502 GTGTCCCGGTTCACTGCCTGTTTCTTGATCTTTGTCAGAGCTATCAGTCATAGCGGGA 561
Qy 47 ----- 47
Db 562 CGGAGGCTGATCAGGTTTCCCGCGACCTGCACGATGGAGTGGGACGCGGCTTC 621
Qy 48 ----- AspAlaCysSerValProVal 54
Db 622 TGTGTTGAGATGCCAGCTGAAGGACGTGGTGTGTTTACAGATGCTGCTCTGTGCTGTA 681
Qy 55 SerGluIleIleAlaValGluGluThrAspValHisGlyHisGlnGlySerGlyLys 74
Db 682 TCTGAGATCATCGCGGTTGAGGAAACAGACGTTTCAACGGGAAACATCAAGGCGTGGAAA 741
Qy 75 TrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArg 94
Db 742 TGGCAGAAATGGAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGCG 801
Qy 95 HisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeu 114
Db 802 CACCGCTGGAAGTGGCGCAGGTGACTTCTGTGTGTCAGAGGAGCAGCTGTGTCACTTG 861
Qy 115 TrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeu 134
Db 862 TGGCTGCAGACCTTCGCGGAGATCTGGAGAAGCTG----- 897
Qy 135 ValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysVal 154
Db 897 ----- 897
Qy 155 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsn 174
Db 898 -----ATTACTGAACATGCTAAT 915
Qy 175 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 194
Db 916 CAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGCGCATCGTGTGTC 975
Qy 195 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSer 214
Db 976 GCGCGAGATGGTATGTTTCAAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGGAGC 1035
Qy 215 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 234
Db 1036 GCGCGGTCGACCAAGAACCCACCCCGGCTGTGCTGGTCCCATAGCTCCGATTTGGA 1095
Qy 235 IleIleProAla-----GlySerThr 241
Db 1096 ATCATTCGCGAGACATGTGGGCCCGGTGAGAACCGCTGTGGCTTGACGCGGTCAACG 1155
Qy 242 AspCysValCysTyrSerThrValClyThrSerAspAlaGluThrSerAlaLeuHistle 261
Db 1156 GACTGCGGTGTGTACTCCACCGTGGCCACGACGACGAGAAACCTCGGCGCTGCATATC 1215
Qy 262 ValVal----- 263
Db 1216 GTTGTGTGCTGTGCCCCGAGGCTCGGAAACCCCGCATCCGCCATCTGCCTGTGGC 1275
Qy 263 ----- 263
Db 1276 AGTGGCACGGGCAGCTGTGCTGGGTGCCAGCTGTGGAACGCATGCTCTGTGAGGCC 1335
Qy 264 -----GlyAspSerLeuAlaMetAspValSer 273
Db 1336 TCGAGGCTTCAGTCCAGGATCGAGAGCCCCGGGACTCGCTGGCCATGGATGTGTCTCA 1395
```

```
Qy 274 ValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyr 293
Db 1396 GTCCACCACAACAGCACACTCTTGCCTACTCCGTGTCCCTGTCTGGGCTACGGCTTCTAC 1455
Qy 294 GlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPhe 313
Db 1456 GGGGACATCATCAAGGACAGTGAAGAAACGGTGGTGGTCTTGGCCACATACGACTTT 1515
Qy 314 SerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuPro 333
Db 1516 TCAGGTATAAAGACCTTCTCTCCACCACCTGCTATGAAGGACAGTGTCTCTCTCCCT 1575
Qy 334 AlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheVal 353
Db 1576 GCACAACACACGCTGGGATCTCCAAGGAATAGGAAGCCCTGCCGGCAGGATGCTTTGTT 1635
Qy 354 CysArgGlnSerLysGlnLeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGlu 373
Db 1636 TGCAGGCAAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAGAAACACTGTATGTTTGGAA 1695
Qy 374 AlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAla 393
Db 1696 GCTGCGGAGGACGTGGAGGAGTGGCAAGTCTCTGTGGGAAGTTTCTGGCCATCAATGCC 1755
Qy 394 ThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeu 413
Db 1756 ACAACATGTCTGTGCTTGTGCGCGGAGCCCCCAGGGGCGCTCTCCCGGCTGCCACTTG 1815
Qy 414 GlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArg 433
Db 1816 GAGAGACGGGTCTTCTGACCTCATCTCCGAAATGCTCCAGGTTCAATTTTCTGAGA 1875
Qy 434 PheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyr 453
Db 1876 TTTCTCATCAGGCACACCAACAGCAGGACAGTTTGACTTCACTTTTGTGAAGTTTAT 1935
Qy 454 ArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLys 473
Db 1936 CGCGTCAAGAAATTCCTGTTTACGTCGAAGCACATGAGGAGTATGAGCAGCGACTCAAG 1995
Qy 474 GluGlyClyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThr 493
Db 1996 GAGGGGGGGAAGAACGCTTTGGGCACATTTGCGAGCAGCCACCCCTCTCTGCTGTCGACC 2055
Qy 494 ValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluVal 513
Db 2056 GTCTCCAAACAGCTCTGGAACCTGGAGCGGAGGTCTCTGCACAGCCCTGCCATCGAGTC 2115
Qy 514 ArgValHisCysGlnLeuValArgLeuPheAlaArgGlyTleGluGluAsnPro 531
Db 2116 AGG-----GCGGTGGTGTATGGCAAGAGGCTGTCTGAAAGGTCCCA 2154
RESULT 12
ACC70838
ID ACC70838 standard; DNA; 3975 BP.
XX
AC ACC70838;
XX
XX 20-NOV-2003 (first entry)
DT
DE Human Sphingosine kinase 4 coding sequence.
XX
XX Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;
XX platelet transfusion; platelet stabiliser; gene; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
CDS 1..1251
FT /*tag= a
ET /product= "Sphingosine kinase 4"
XX
XX W02003031627-A1.
```

XX 17-APR-2003.
XX 28-SEP-2001; 2001WO-JP008537.
XX 28-SEP-2001; 2001WO-JP008537.
XX (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.
XX (CHBI-) CHEM BIOLOGY INST.
XX Igarashi Y, Kihara A;
XX WPI: 2003-354917/33.
XX P-PSDB; ABR56301.
XX Platelet derived polypeptides with sphingosine kinase activity for
XX treatment of sphingosine related disorders.
XX Claim 4; Page 30-31; 39pp; Japanese.
XX
XX The present sequence is the coding sequence for human sphingosine kinase
XX 4 (SPHK4). The kinase can be used for the diagnosis and treatment of
XX sphingosine related disorders. The kinase can also be potentially used
XX for controlling toxicity of platelet transfusion and as a platelet
XX stabiliser
XX
SQ Sequence 3975 BP; 943 A; 978 C; 1042 G; 1012 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3 87e-192 Length: 3975
Score: 2210.00 Matches: 416
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.52% Indels: 0
DB: 8 Gaps: 0

US-10-631-958-10 (1-537) x ACC70838 (1-3975)

QY 122 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 141
DB 1 ATGCTGGAGAGCTGACGTCACAGCAAAAGCATTTTACTGGTATTATCAACCCGTTTGA 60

QY 142 GlyLysGlyGlnGlyArgIleThrGluArgLysValAlaProLeuPheThrLeuAla 161
DB 61 GGAAAGGACACAGGACGGATATGAAAGAAAGTGGCACCACCTGTTACCTTAGCC 120

QY 162 SerIleThrThrAspIleThrValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyr 181
DB 121 TCCATCACACATGATCATCTGTTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTAT 180

QY 182 GluIleAsnIleAspLysThrAspGlyIleValCysValGlyGlyAspGlyMetPheSer 201
DB 181 GAGATTAACATAGACAAATACGACGCGCATCGTCTGTGCGGGAGATGTTATGTTACG 240

QY 202 GluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHis 221
DB 241 GAGGTGCTGACGGTCTGATTGGGAGGACGACAGAGGCGCGGGGTCCAGCAGAACCC 300

QY 222 ProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThr 241
DB 301 CCCCAGGCTGTCTGCTGCTCCCATAGTACCTCCGGAATTGGAATCATTTCCCGCAGGGTCAACG 360

QY 242 AspCysValCysThrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIle 261
DB 361 GACTGGGTGTGTACTCCACCGTGGGACCAGCGAGCGAGAAACCTCGCGCTGCATATC 420

QY 262 ValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeu 281
DB 421 GTTGTGGGGACTCGCTGGCCATGGATGTGCTCCTCAGTCCACCAACAGCAGACACTCCTT 480

QY 282 ArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGlu 301
DB 481 CGCTACTCGGTCTGCTGGGCTACGGCTTCTACGGGGGACATCATCAAGGACAGTGAG 540

QY 302 LysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSer 321
DB 541 AAGAAACGGTGGTGGTCTTGCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCC 600

QY 322 HisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerPro 341
DB 601 CACCACCTGCTATGAAGGACAGTGTCTCTCTCCGACCAACACAGCGTGGGATCTCCA 660

QY 342 ArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeu 361
DB 661 AGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGCAGGCAAGCAACAGCAGCTG 720

QY 362 GluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaGluAspValGluGluTrp 381
DB 721 GAGGAGGACGAGAAAGACACTGTATGTTTGAAGCTGCGAGGAGCTGCGAGGAGTGG 780

QY 382 GlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArg 401
DB 781 CAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCCAACATGTCTCTGCTGTTCGC 840

QY 402 ArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIle 421
DB 841 CGGAGCCCGGAGGCGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATC 900

QY 422 LeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGln 441
DB 901 CTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACCAACCCAG 960

QY 442 GlnAspGlnPheAspPheThrPheValIleValTyrArgValLysLysPheGlnPheThr 461
DB 961 CAGGACCAAGTTGACTTTCACCTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTTACG 1020

QY 462 SerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPheGly 481
DB 1021 TCGAAGCACATGGAGGATGAGCAGCAGCTCAAGGAGGGGGGAGAGCGCTTTGGG 1080

QY 482 HisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCys 501
DB 1081 CACATTTGCAGCAGCAGCCCTCTCTGCTGCTGCACCCGTCTCCAACAGCTCTCGAACTGC 1140

QY 502 AspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArg 521
DB 1141 GACGGGAGGTCTCTGCACAGCCCTGCCATCAGAGTCAGAGTCCACTCCAGCTGGTTGCA 1200

QY 522 LeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
DB 1201 CTCCTTGCACGAGGAATTGAGAGAAATCCGAGGACAGACTCACACAGC 1248

RESULT 13
AAS77730
ID AAS77730 standard; cDNA; 2186 BP.
XX
AC AAS77730;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13534.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX
XX WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG13543.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 1; SEQ ID NO 13534; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activities. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 2186 BP; 455 A; 637 C; 669 G; 425 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 5,14e-147 | Length: | 2186 |
| Score: | 1714.50 | Matches: | 345 |
| Percent Similarity: | 63.65% | Conservative: | 0 |
| Best Local Similarity: | 63.65% | Mismatches: | 0 |
| Query Match: | 59.37% | Indels: | 197 |
| DB: | 5 | Gaps: | 3 |

US-10-631-958-10 (1-537) x AAS77730 (1-2186)

| | | | |
|----|-----|---------------------------------------------------------------|-----|
| Qy | 170 | ThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAsp | 189 |
| Db | 3 | ACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGAC | 62 |
| Qy | 190 | Gly----- | 190 |
| Db | 63 | GGATGCTACTGTCCTTCAAGGATGCCCTGTCTCTCGAAACCACAGCCCGCTGGGAAG | 122 |
| Qy | 190 | ----- | 190 |
| Db | 123 | GCTGCCCGGTGCAAGGTGTAGGTACGCGGAAGGGGMAAGAGCCTTCCCATGGAGCCG | 182 |
| Qy | 190 | ----- | 190 |
| Db | 183 | GGCCCAAGACAGATGCTCTGCCAAGGGCCAGGTGTGCAGGCTGCCCTCTCTGAAGCTGAA | 242 |
| Qy | 190 | ----- | 190 |
| Db | 243 | GCCTCAGGCCCTGCTCGCTCAGAGGGCGCACCTGCAGGAGCCCTGCACCGCTGTGTTC | 302 |
| Qy | 191 | -----IleValCysValGlyGlyAspGlyMetPheSerGluValIleHisGlyLeu | 207 |
| Db | 303 | TCCTGCAGCATCGTCTGTGTCGGGGAGATGGTATGTCACGAGGTGCTGCACGGTCTG | 362 |
| Qy | 208 | IleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuVal | 227 |

| | | | |
|----|------|---------------------------------------------------------------|------|
| Db | 363 | ATTGGGAGGACGACAGAGGCGCGGGTCCAGACAGAACACCCCGGGGTGTCTGGTC | 422 |
| Qy | 228 | ProSerSerLeuArgIleGlyIleIleProAla----- | 238 |
| Db | 423 | CCAGTAGCCTCCGGATTGGAATCATTTCCCGCAGGCTTTGCAAAATGACCTGGCAGGGAC | 482 |
| Qy | 238 | ----- | 238 |
| Db | 483 | GAGGTGTCTCTCTCTCTGCGCCCTGTGTCTGGCCCCGAGGGTGGCGGCATGTGTGCACACT | 542 |
| Qy | 238 | ----- | 238 |
| Db | 543 | TTCACTCTCGTACAGCTCTGGGATGTAGCACCGCAGTCATCCCAATTTTATGGATGAA | 602 |
| Qy | 239 | -----GlySer | 240 |
| Db | 603 | GACAGAGGACTCGGGAGCATGTGGGCCCGCGGTGAGAACCGTGGTGGCTTGGACGGGTCA | 662 |
| Qy | 241 | ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis | 260 |
| Db | 663 | ACGAGCTGCGTGTGTACTCCACCGTGGGCACCGACGACGACAGAACTCTGGCGCTGCAT | 722 |
| Qy | 261 | IleValVal----- | 263 |
| Db | 723 | ATCGTTGTGGTCTGCTGCCCGAGGCTCGGAAACACCCCGCTCCCGCCATACTGCCTGT | 782 |
| Qy | 263 | ----- | 263 |
| Db | 783 | GGCAGTGGGCACGGGACGTGTGCTGGGTGCGCAGTGTGGAAACGCGTCTGTGTGAG | 842 |
| Qy | 264 | -----GlyAspSerLeuAlaMetAspValSer | 272 |
| Db | 843 | GCCTCGAGGCTTCAGTCCAGATGCAGAGCCCGGGGACTCGCTGGCCATGGATGTCTCC | 902 |
| Qy | 273 | SerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe | 292 |
| Db | 903 | TCAGTCCACCAACACAGCACACTCTTCCTGCTACTCCCTGCTCCCTGCTGGGTACGGCTTC | 962 |
| Qy | 293 | TyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAsp | 312 |
| Db | 963 | TACGGGACATCATCAAGGACAGTGAAGAAACCGGTGGTGTGGTCTTTCAGATACGAC | 1022 |
| Qy | 313 | PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu | 332 |
| Db | 1023 | TTTTCAGTTTAAAGACTTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTCTCCTC | 1082 |
| Qy | 333 | ProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe | 352 |
| Db | 1083 | CCTGCACAAACACACGCTGGGATCTCCAAGGGATAGGAAGCCCTGCCGGCAGGATGCTTT | 1142 |
| Qy | 353 | ValCysArgGlnSerLysGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeu | 372 |
| Db | 1143 | GTTTGCAGGCAAGCAAGCAGCAGCTGGAGAGAGTGCAGTGTGTGGGAAGTCTGTATGGTTG | 1202 |
| Qy | 373 | GluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsn | 392 |
| Db | 1203 | GAAGTCCGGAGGACGCTGGAGAGTGCAGTGTGTGGGAAGTCTTCTGGCCATCAAT | 1262 |
| Qy | 393 | AlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHis | 412 |
| Db | 1263 | GCCACAAACATGCTCTGTGCTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCAC | 1322 |
| Qy | 413 | LeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeu | 432 |
| Db | 1323 | TTGGAGACGGGCTCTCTGACCTCATCTCCGAAATGCTCCAGGTTCAATTTTCTG | 1382 |
| Qy | 433 | ArgPheLeuIleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluVal | 452 |
| Db | 1383 | AGATTCTCATCAGGCACACCAACCCAGCAGACCCAGTTTCACTTTCTTTTGTGAAGTT | 1442 |
| Qy | 453 | TyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu | 472 |

Db 1443 TATCCGCTCAAGAAATTCAGTTTACGTGCGAAGCAATGAGGATGAGGACAGCGACCTC 1502
 Qy 473 LysGluGlyGlyLysArgPheGlyHisIleCysSerSerHisProSerCysCys 492
 Db 1503 AAGGAGGGGGGAGAGAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCTGCTGCTGC 1562
 Qy 493 ThrValSerAsnSerSerTyrAsnCysAspGlyGluValLeuHisSerProAlaIleGlu 512
 Db 1563 ACCGCTCTCAACAGCTCTCTGGAACCTGCACGGGAGGTCTCTGCAGCCCTGCCATCGAG 1622
 Qy 513 ValArg 514
 Db 1623 GTCAGA 1628
 RESULT 14
 ID ABL40822
 XX ABL40822 standard; cDNA; 979 BP.
 AC ABL40822;
 XX
 DT 03-JUL-2002 (first entry)
 DE Human sphingosine kinase-like protein encoding cDNA.
 XX
 KW Human sphingosine kinase-like protein; intracellular signalling; gene;
 KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
 KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 2..789
 FT /*tag= a
 FT /product= "sphingosine kinase-like protein"
 FT /note= "start and stop codons are not indicated"
 XX
 PN WO200228906-A2.
 XX
 XX
 PD 11-APR-2002.
 XX
 XX 05-OCT-2001; 2001WO-EP011516.
 XX
 PR 06-OCT-2000; 2000US-0238005P.
 PR 23-AUG-2001; 2001US-0314113P.
 XX
 PA (FARB) BAYER AG.
 XX
 XX Kossida S, Encinas J;
 XX
 XX WPI; 2002-340094/37.
 DR P-PSDB; ABB07854.
 XX
 XX New reagent for modulating the activity of sphingosine kinase-like
 PT protein polypeptide or polynucleotide and treating cancer, asthma,
 PT allergy, an autoimmune disease, or a central or peripheral nervous system
 PT disorder.
 XX
 XX Claim 1; Fig 1; 120pp; English.
 PS
 XX The invention relates to a human sphingosine kinase-like protein. The
 CC polypeptide can be expressed by standard recombinant methodology. The
 CC sphingosine kinase-like protein and gene can be used to regulate
 CC intracellular signalling and consequently cell proliferation and
 CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
 CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
 CC peripheral nervous system disorders (e.g. Parkinson's disease). The
 CC present sequence represents the human sphingosine kinase-like protein
 CC encoding cDNA
 XX
 SQ Sequence 979 BP; 237 A; 259 C; 271 G; 212 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.04e-140 Length: 979

Score: 1640.50 Matches: 314
 Percent Similarity: 96.63% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 56.80% Indels: 11
 DB: 6 Gaps: 1
 US-10-631-958-10 (1-537) x ABL40822 (1-979)
 Qy 130 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 149
 Db 2 CCRAAGCATTTTACTTGTATTTTATCAACCCGTTTGGAGAGAAAGGACAAGCAACGCGATA 61
 Qy 150 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIle--- 168
 Db 62 TATGAAAGAAAGTGGCACCACCTGTTCCACCTTAGCCCTCCATCACCACCTGATCATCGGT 121
 Qy 169 -----ValThrGluHisAlaAsnGlnAlaLysGlu 178
 Db 122 AACAAATTCCTATGTTAACTATGTAGAGTAATTTACTGAAACATGCTAATCAGGCCAAGAG 181
 Qy 179 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 198
 Db 182 ACTCTGTATGAGATTAAACATAGACAATACAGCGCATCGTCTGTGTCCGCGAGATGGT 241
 Qy 199 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 218
 Db 242 ATGTTTCAGCGAGGTGCTGCACGGTCTGATTTGGGAGGACGAGAGCGCGGGGTGCAC 301
 Qy 219 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 238
 Db 302 CAGAACCAACCCCGGGCTGTGGTCTCCAGTAGCCCTCCGGATTGGAATCAITCCCGCA 361
 Qy 239 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 258
 Db 362 GGGTCAACGAGCTCGGTGTCTTACTCCACCGTGGCACACGAGCAGCAGAACCTCGCG 421
 Qy 259 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 278
 Db 422 CTGCATATCGTTGTTGGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACACAGC 481
 Qy 279 ThrLeuLeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIleIleIle 298
 Db 482 ACACCTCTCGTACTCCGTGTCCCTGCTGGGTACGGCTTCTACGGGGACATCATCAAG 541
 Qy 299 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 318
 Db 542 GACAGTGAGAAGAAACGGTGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGACC 601
 Qy 319 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 338
 Db 602 TTCCTCTCCACCACTGCTATGAAGGGACAGTGTCTTCTCCCTGACACACACACGCTG 661
 Qy 339 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 358
 Db 662 GGATCTCCAAGGATAGAAAGCCCTGCGGGGAGAGATGCTTTGTTTCAGGCAAGCAAG 721
 Qy 359 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 378
 Db 722 CAGCAGCTGGAGGGAGGAGCAGAGAAGCACTGTATGGTTTGGAGCTGCGGAGGAGCTG 781
 Qy 379 GlnGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 398
 Db 782 GAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCACAAAACATGCTCCTGT 841
 Qy 399 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 418
 Db 842 GCTGTGTCGGGAGGCCCCAGGGGCTCTCCCGGCTGCCACCTTGGGAGAGCGGCTCTTCT 901
 Qy 419 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 438
 Db 902 GACTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTTCTCATCAGGCAC 961
 Qy 439 ThrAsnGlnGlnAspGln 444

Qy 516 HisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnPro 531
Db 1448 -----GCCGTGGTGATGGCAAGAGGCTGCGAAGGTCCCA 1483

Search completed: September 6, 2005, 11:25:27
Job time : 893.053 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 5, 2005, 23:41:54 ; Search time 1512.64 Seconds
(without alignments)
2324.913 Million cell updates/sec

Title: US-10-631-958-10
Perfect score: 2888
Sequence: 1 NGATGAEPLOSVLWKQOR.....QLVRLFARGIENPKPDHS 537

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10631958/runat_02092005_165813_4470/app.query.fasta_1.1941
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOFCPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALTN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10631958 @CGN 1 1 1588 @runat_02092005_165813_4470
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------------------|-------------------|
| 1 | 2888 | 100.0 | 1614 | 10 | US-09-969-896-9 | Sequence 9, Appli |
| 2 | 2888 | 100.0 | 1614 | 19 | US-10-631-958-9 | Sequence 9, Appli |
| 3 | 2888 | 100.0 | 1740 | 18 | US-10-262-511-39 | Sequence 39, Appl |
| 4 | 2888 | 100.0 | 4413 | 10 | US-09-969-896-16 | Sequence 16, Appl |
| 5 | 2888 | 100.0 | 4413 | 19 | US-10-631-958-16 | Sequence 16, Appl |
| 6 | 2888 | 100.0 | 4429 | 19 | US-10-618-941-55 | Sequence 55, Appl |
| 7 | 2881 | 99.8 | 4432 | 17 | US-10-120-988-148 | Sequence 148, App |
| 8 | 2880 | 99.7 | 4463 | 16 | US-10-315-597A-1 | Sequence 1, Appli |
| 9 | 2649.5 | 91.7 | 1840 | 9 | US-09-784-810A-5 | Sequence 5, Appli |
| 10 | 2649.5 | 91.7 | 1840 | 22 | US-10-876-281-5 | Sequence 5, Appli |
| 11 | 1640.5 | 56.8 | 979 | 10 | US-09-969-896-1 | Sequence 1, Appli |
| 12 | 1640.5 | 56.8 | 979 | 19 | US-10-631-958-1 | Sequence 1, Appli |
| 13 | 1032 | 35.7 | 817 | 19 | US-10-115-635-247 | Sequence 247, App |
| 14 | 871 | 30.2 | 550 | 10 | US-09-969-896-8 | Sequence 8, Appli |
| 15 | 871 | 30.2 | 550 | 19 | US-10-631-958-8 | Sequence 8, Appli |
| 16 | 811 | 28.1 | 474 | 10 | US-09-969-896-4 | Sequence 4, Appli |
| 17 | 811 | 28.1 | 474 | 19 | US-10-631-958-4 | Sequence 4, Appli |
| C 18 | 656 | 22.7 | 522 | 9 | US-09-784-810A-7 | Sequence 7, Appli |
| C 19 | 616.5 | 21.3 | 2657 | 20 | US-10-876-281-7 | Sequence 7, Appli |
| 20 | 616.5 | 21.3 | 2657 | 18 | US-10-425-115-176413 | Sequence 176413, |
| 21 | 573.5 | 19.9 | 2079 | 18 | US-10-425-114-34530 | Sequence 34530, A |
| 22 | 573.5 | 19.9 | 2079 | 20 | US-10-425-115-86211 | Sequence 86211, A |
| 23 | 525 | 18.2 | 329 | 10 | US-09-969-896-5 | Sequence 5, Appli |
| 24 | 525 | 18.2 | 329 | 19 | US-10-631-958-5 | Sequence 5, Appli |
| C 25 | 480.5 | 16.6 | 2189 | 19 | US-10-437-963-92801 | Sequence 92801, A |
| 26 | 437.5 | 15.1 | 564 | 22 | US-10-477-445-46 | Sequence 46, Appl |
| 27 | 399.5 | 13.8 | 1833 | 19 | US-10-437-963-30650 | Sequence 30650, A |
| 28 | 389 | 13.5 | 2084 | 18 | US-10-424-599-47396 | Sequence 47396, A |
| 29 | 376 | 13.0 | 2166 | 20 | US-10-425-115-64890 | Sequence 64890, A |
| 30 | 371.5 | 12.9 | 1869 | 18 | US-10-425-114-34806 | Sequence 34806, A |
| 31 | 371 | 12.8 | 339 | 9 | US-09-783-590-5271 | Sequence 5271, Ap |
| 32 | 363 | 12.6 | 3362 | 20 | US-10-425-115-29002 | Sequence 29002, A |
| 33 | 347 | 12.0 | 2014 | 18 | US-10-424-599-59419 | Sequence 59419, A |
| 34 | 344.5 | 11.9 | 1512 | 20 | US-10-425-115-64892 | Sequence 64892, A |
| 35 | 343.5 | 11.9 | 382 | 9 | US-09-784-810A-9 | Sequence 9, Appli |
| 36 | 343.5 | 11.9 | 382 | 22 | US-10-876-281-9 | Sequence 34081, A |
| 37 | 339 | 11.7 | 1774 | 18 | US-10-425-114-34081 | Sequence 95148, A |
| 38 | 331.5 | 11.5 | 1386 | 19 | US-10-437-963-95148 | Sequence 77896, A |
| 39 | 330.5 | 11.4 | 3008 | 19 | US-10-437-963-77896 | Sequence 32139, A |
| 40 | 326.5 | 11.3 | 1857 | 20 | US-10-425-114-32139 | Sequence 97849, A |
| 41 | 326.5 | 11.3 | 2247 | 18 | US-10-425-115-97849 | Sequence 58701, A |
| 42 | 322 | 11.1 | 1916 | 19 | US-10-437-963-58701 | Sequence 25, Appl |
| 43 | 319 | 11.0 | 2609 | 17 | US-10-348-052-25 | Sequence 25, Appl |
| 44 | 319 | 11.0 | 2609 | 19 | US-10-622-011-25 | Sequence 79424, A |
| 45 | 316.5 | 11.0 | 2284 | 20 | US-10-425-115-79424 | |

ALIGNMENTS

RESULT 1
US-09-969-896-9
; Sequence 9, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-9

Alignment Scores:
Pred No.: 0 Length: 1614
Score: 2888.00 Matches: 537
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-631-958-10 (1-537) x US-09-969-896-9 (1-1614)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20
Db 1 ATGGGGCGACGGGGCGGGCGAGCCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC 60

Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 61 TGGCCCGTGAGCCCTGGAGCCCGCGGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 120

Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db 121 GGAGCCGGCGCCCGGGCGGGATGCCCTGCTCTGTGCTGTATCTGATCATCCCGTT 180

Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 181 GAGGAAACAGACGTTTACGGGAAACATCAAGGCAGTGGAAATGGCAGAAATGGAAAG 240

Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100
Db 241 CCTTACGCTTTTACAGTTTCACTGTGTATAAGAGAGACGACGGCACCGCTGGAACTGGCG 300

Qy 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db 301 CAGTGACTTCTGGTGTCCAGAGAGAGCAGCTGTGTCTCTGTGGTGGCGACCTCGGG 360

Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 361 GAGATGCTGGAGAAAGCTGACGCTCCAGACCAAGCAATTTACTGGTATTATCAACCCGTTT 420

Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 421 GGAGGAAAGGACAAAGGCAAGCGGATATATGAAGAAAGTGGCACCACCTGTTTACCTTA 480

Qy 161 AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 481 GCCTCCATCACCACCTGACATCATGTTACTGAAATGCTATATCAGGCCAAGGACTCTG 540

Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 541 TATGAGATTAACATAGACAAATACGCGCATCGTCTGTGTGGCGGAGATGTTATGTTTC 600

Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 601 AGCGAGGTGCTGCAACGCTGCTATTTGGGAGGACGAGAGAGCGCGCGGGTTCGACGAAAC 660

Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 661 CACCCCGGGCTGTGCTGGTCCCGCAGTAGCTCCGATTGGAATCATTTCCCGCAGGGTCA 720

Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 721 ACGGACTCGGTGTGTTACTCCACCGTGGGCACCGACGACGAGAAACCTCGGCGCTGCAT 780

Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrIleu 280
Db 781 ATCGTTGTGGGACTCGCTGGCCATGATGTGCTCCTCAGTCCACCAACAGCACACTC 840

281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
841 CTTTGGCTACTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGCAGT 900

301 GluLysLysArgTrpTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
901 GAGAAGAAACGGTGGTGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGACCTCTCCTC 960

321 SerHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
961 TCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCGCAACACACGTTGGGATCT 1020

341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
1021 CCAAGGGATAGGAAGCCCTCGCGGAGGATGCTTTGTTTGCAGGCCAAAGCAGCAG 1080

361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
1081 CTGGAGGAGGAGCAGAAAGACATGTATGTGTTTGGAACTCGGAGGACGTGGAGGAG 1140

381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
1141 TGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCCAACATGTCTGTGCTGT 1200

401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
1201 GCGCGAGGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1260

421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
1261 ATCTCATCTCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACCAAC 1320

441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
1321 CAGCAGACCGAGTTGACTTCACTTTTGTGTAAGTTTATCCGTCAGAAATTTCCAGTTT 1380

461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
1381 ACGTGGAAGCACATGGAGGATGAGGACAGCGACTCAAGGAGGGGGGGAAGCGCTTT 1440

481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
1441 GGGCACATTTTGCAGACGCCACCCCTCTCTGTCTGTGTCACCGCTCCCAACAGCTCTGGAAC 1500

501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
1501 TGGACGGGGAGGTCTTCACAGCCCTGCCATCGAGTCCAGTCCACTGCCAGCTGGTT 1560

521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
1561 CGACTCTTTGCAGGAAATTGAAGAAATTCGAAGACGAGCTCACACAGC 1611

RESULT 2
US-10-631-958-9
; Sequence 9, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631.958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1614
```



```
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqlist version 0.1
; SEQ ID NO 39
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (76)..(1686)
US-10-262-511-39

Alignment Scores:
Pred. No.: 0 Length: 1740
Score: 2888.00 Matches: 537
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-631-958-10 (1-537) x US-10-262-511-39 (1-1740)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20
Db 76 ATGGGGCGACGGGGCGGGCGGACCGCTGCAATCCGTGCTGGTGAAGCAGCAGCGC 135

Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 136 TGCGCCGTGAGCTGGAGCCCGCGGCTCTGCTGCGTGTGGTGGAGCCCGGGGCC 195

Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db 196 GGAGCCGGCGCCCGCGTGTGATGCCCTGCTGTGCTGTATCTGAGATCATCGCCGTT 255

Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 256 GAGGAACACAGACGTTACGGGAACATCAAGCAGTGGAAATGCGAGAAATGGAAAG 315

Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgHisArgTrpLysTrpAla 100

316 CCTTACGCTTTTACAGTTCACTGCTAAAGAGACGACGACGCGCTGGGCG 375
Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnTrpLeuArg 120
Db 376 CAGGTGACTTTCGTGGTCCAGAGGAGCAGCTGTGTCTACCTTGTGGCTGCAGACCTCGGG 435
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 436 GAGATGCTGGAGAAGCTGACGCTCCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTT 495
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 496 GGAGGAAAAGGACAAAGCGATATATGAAAGAAAAGTGGCACCACCTGTTCCACCTTA 555
Qy 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 556 GCCTCCATCACCACCTGACATCATCGTTACTGAACATGCTAATCAGGCCAAGGAGACTCTG 615
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 616 TATGAGATTAAACATAGACAAATACGACGGCATCGTGTGTCTGGCGGAGATGGTATGTT 675
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 676 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGTGCAGCAGAAC 735
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 736 CACCCCGGGCTGTGCTGGTCCCAGTAGCTCCGGATTGGAATCATTCCTCCGAGGGTCA 795
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 796 ACGGACTGCGTGTGTTACTCCACCGTGGGCACCGACGACGAGAAACCTCGGCGTGCAT 855
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSerThrLeu 280
Db 856 ATCGTTGTTGGGAGCTCGCTGGCCATGGATGTGCTCTCAGTCCACCACACAGCACACTC 915
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
Db 916 CTTGCGTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGT 975
Qy 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
Db 976 GAGAAGAAACGGTGGTGGTGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCCT 1035
Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db 1036 TCCACCACTGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACACGCTGGGATCT 1095
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db 1096 CCAAGGGATAGGAAGCCCTCCCGGAGGATGCTTTGTTGCGAGGCAACAGACAGCAG 1155
Qy 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db 1156 CTGGAGGAGGAGCAGAGAAGACACTGTATGGTTTGGAACTGCGGAGGACGTTGGAGGAG 1215
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db 1216 TGGCAAGTCGCTGTGGAAAGTTTCTGGCCATCAATGCCCAACACATGCTCTGTGCTGT 1275
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db 1276 CGCGGAGCCCCAGGGCCCTCTCCCGGCTGCCACCTTGGGAGACGGGTCTTCTGACCTC 1335
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 1336 ATCCTCATCCGAAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAAC 1395
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
```

Db 1396 CAGCAGGACAGTTTTCACCTTTTGTGTAAGTTTATCGCGTCAAGAAATTCAGTTT 1455

Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480

Db 1456 ACGTCGAAGCACATGAGGATGAGGACAGCAGCTCAAGAGCGGGGAAGAAGCGCTTT 1515

Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500

Db 1516 GGGCAATTTGAGCAGGACCCCTCTGCTGCTGACCGTCTCAACAGAGCTCCCTGGAAC 1575

Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520

Db 1576 TCGCAGCGGGAGGTCTCCACAGCCCTGCCATCGAGGTGAGAGTCCATGCCAGCTGGTT 1635

Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537

Db 1636 CGACTCTTTGACGAGGAATTGAAGAAATCCGAAGCCAGACTCACAGC 1686

RESULT 4

US-09-969-896-16

; Sequence 16, Application US/09969896

; Publication No. US20030125533A1

GENERAL INFORMATION:

; APPLICANT: Kossida, Sophia

; TITLE OF INVENTION: Regulation of human Sphingosine

; TITLE OF INVENTION: Kinase-Like Protein

; FILE REFERENCE: 004974.00594

; CURRENT APPLICATION NUMBER: US/09/969,896

; CURRENT FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: US 60/238,005

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: US 60/314,113

; PRIOR FILING DATE: 2001-08-23

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 4413

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-969-896-16

Alignment Scores:

Pred. No.: 0 Length: 4413

Score: 2888.00 Matches: 537

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 10 Gaps: 0

US-10-631-958-10 (1-537) x US-09-969-896-16 (1-4413)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20

Db 76 ATGGGGGCGACGGGGGGCGGAGCGCGTGCATCCGTGCTGTGGGTGAGCAGCAGCGC 135

Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40

Db 136 TCGCGCGTGAGCTGAGAGCCCGCGGGCTCTGCTGCGCTGTGGCGGAGCCCGGGGCC 195

Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60

Db 196 GGAGCGGGCGCCCGCGCGGATGCTGCTCTGTGCGCTGTATCTGAGATCATCGCGGT 255

Qy 61 GluGluThrAspValHisGlyLysHisGlnGlnSerGlyLysTrpGlnLysMetGluLys 80

Db 256 GAGGAACACAGAGCTTCACGGGAAACATCAAGCAGTGGAAATGCGAAGATGGAAG 315

Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100

Db 316 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGCCTGGAAAGTGGCG 375

Qy 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120

Db 376 CAGGTGACTTTCTGGTGTCAGAGGAGCAGCTGTGTCTCACTTGTGGCTGCGAGCCCTGCGG 435

Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140

Db 436 GAGATGCTGGAGAAGCTGACGCTCCAGACCAAGCAATTTTACTGGTATTATCAACCCGTTT 495

Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160

Db 496 GGAGGAAAGACCAAGCGGATATATGAAAGAAAGTGGCACCACTGTGTTCACCTTA 555

Qy 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180

Db 556 GCCTTCATCACACTGACATCATCTTACTGAAACATGCTATATCAGGCCACGAGACTCTG 615

Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200

Db 616 TATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTCTGCGCGGAGATGGTATGTT 675

Qy 201 SerGluValLeuHisGlyLysGlyIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220

Db 676 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGAGCGCGGGTTCGACCAGAAC 735

Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240

Db 736 CACCCCGGGCTGTGCTGCCCATGAGTCCCGATTCGATTCCTCCCGCAGGGTCA 795

Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260

Db 796 ACGGACTGCGTGTGTTACTCCACCGTGGGACCAGCAGCAGCAAACTCTCGCGCTGCAT 855

Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280

Db 856 ATCGTTGTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACACACAGCACACATC 915

Qy 281 LeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300

Db 916 CTTGCTACTCCGTGCTCCGCTGCGCTACCGCTTCTACGGGACATCATCAAGACAGT 975

Qy 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320

Db 976 GAGAAGAAACCGTGGTGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC 1035

Qy 321 SerHisHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340

Db 1036 TCCACCACTCTATGAAGGACAGTGTCTCTCCTCGCACACACACACGCTGGGATCT 1095

Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360

Db 1096 CCAAGGATAGAAAGCCCTGCGGGCAGATGCTTTGTTTCAGGCAAGCAAGCAGCAG 1155

Qy 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380

Db 1156 CTGGAGGAGGAGCAGAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGCAGCTGGAGGAG 1215

Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAlaThrAsnMetSerCysAlaCys 400

Db 1216 TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGTCTGT 1275

Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420

Db 1276 CGCCGAGCCCCCAGGGGCTCTCCCGGCTGCCCATTTGGGAGACGGGTCTTCTGACCTC 1335

Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440

Db 1336 ATCTCTATCCGAAATGCTCCAGGTTCAAATTTTCTGAGATTTCTCATCAGCACACCAAC 1395

Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460

Db 1396 CAGCAGGACCAAGTTTGCATTCACCTTTTGTGAAGTTTATTCGCGTCAAGAAATTCAGTTT 1455

Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480

Db 1456 ACGTCGAAGCACATGGAGGATGAGGACAGCAGCTCAAGGAGGGGGGAAGAGCGCTTT 1515

```
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
Db 1516 GGGACATTTGCAGACGCCACCTCTCTGCTGCTGCACCGTCTCCACAGCTCTCTGGAAC 1575
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 1576 TGCAGCGGGAGGTCCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCCAGCTGGTT 1635
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
Db 1636 CGACTCTTTGCAGGAGGAATTAAGAGAATCCGAAGCCAGACTCACACAGC 1686

RESULT 5
US-10-631-958-16
; Sequence 16, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631.958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 4413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-16

Alignment Scores:
Pred. No.: 0 Length: 4413
Score: 2888.00 Matches: 537
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-631-958-10 (1-537) x US-10-631-958-16 (1-4413)
Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
Db 76 ATGGGGCGACGGGGCGGGAGCGCGTGCATCCGTGCTGTGGTGAACGACGACGC 135
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuArgTrpTrpArgSerProGlyPro 40
Db 136 TCGCGCGTGAGCTGGAGCCCGCGGGCTCTGCTGCGCTGGTGGCGGAGCCCGGGGCC 195
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db 196 GGAGCGCGCGCGCGCGCGGATGCCCTGCTGTGCCCTGTATCTGAGATCATCGCGTT 255
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 256 GAGGAACACAGACGTTACGGGAACATCAAGCAGTGGAAAATGGCAAAATGGAAAAG 315
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100
Db 316 CTTTACCGCTTTTACAGTTTCACTGTGTAAAGAGAGCACGACGCCACCGCTGGAAGTGGCG 375
Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db 376 CAGGTGACTTTCTGTGTGTCAGAGAGCAGCTGTGTCTCTGTGCTGCACACCTTCGG 435
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
```

```
Db 436 GAGATGCTGAGAAGCTGACGCTCCAGACCAAAACATTTACTGGTATTATCAACCCGTTT 495
Qy 141 GlyGlyLysGlyGlnGlyLysArgGlyIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 496 GAGGAAAGACCAAGCAAGCGGATATATGAAGAAAAAGTGGCACCACCTGTTTCCACTTA 555
Qy 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 556 GCCTCCATCAACCTGACATCATCTGTTACTGACATGCTAATCAGGCCAAGGAGACTCTG 615
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 616 TATGAGATTAACTAGACAAATACGACGCGCATCGTCTGTCTGCGCGAGATGGTATGTTTC 675
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 676 AGCGAGGTGCTGCACGGTCTGANTTGGGAGGACGACAGAGGAGCGCGGGTGCACCAAC 735
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 736 CACCCCGGGCTGTGCTGGTCCCGAGTAGCTCCGATTCGATTCATTCCTCCGACAGGTCA 795
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 796 ACGGACTGCGTGTGTTACTCCACCGTGGCACACGACGACGAGAAACCTCGGCGCTGCAT 855
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
Db 856 ATCGTGTGTTGGGGACTCGCTGGCCATGGATGTGCTCCTCAGTCCACCAACACAGCAGACTC 915
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
Db 916 CTTGCTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGACAGT 975
Qy 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
Db 976 GAGAAGAAACGGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTTAAAGACCTTCCTC 1035
Qy 321 SerHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db 1036 TCCACACACTGCTATGAAGGAGACAGTGTCTCTCCCTCGCACCAACACAGCGTGGATCT 1095
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db 1096 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTCTTTGTCAGGCAAGCAACAGCAGAG 1155
Qy 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db 1156 CTGGAGGAGGAGCAGAAAGACACTGTATGGTTTGAAGCTGCGGAGGAGCTGGAGGAG 1215
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db 1216 TGGCAAGTCGTGTGGGAAGTTTCTGGCCATCAATGCCCAAAACATGTCTCTGTGCTGT 1275
Qy 401 ArgArgSerProArgGlyLeuSerProAlaHisLeuGlyAspGlySerSerAspLeu 420
Db 1276 CGCGGAGACCCAGGGGCGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTCTTGACCTC 1335
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 1336 ATCTCATCCGGAATGCTCCAGGTTCAAATTTCTGAGATTTTCTCATCAGGCACCAAC 1395
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysValPheGlnPhe 460
Db 1396 CAGCAGACCATGTTGACTTTCACCTTTTGTGAAGTTTATCGCGTCAGAAATTCAGTTT 1455
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
Db 1456 ACGTCTGAAGCACATGGAGGATGAGCAGCGACCTCAAGGAGGGGGGGAAGACGCTTT 1515
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
```

Db 1516 GGGCACANTTTCAGCAGCCACCCCTCTGCTGTGCACCGGTCTCCAACAGCTCTCTGGAAC 1575

Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 1576 TGCAGCGGGAGGTTCTTGCAAGCCCTGCCATCGAGGTCAAGTCCACTGCCAGTGTGTT 1635

Qy 521 ArgLeuPheAlaArgGlyVileGluGluAsnProLysProlepSerHisSer 537
Db 1636 CGACTCTTTGCACGAGGANTTTGAAGAAGATCCGAAGCCAGACTCACACAGC 1686

RESULT 6

US-10-618-941-55

; Sequence 55, Application US/10618941

; Publication No. US20040197792A1

; GENERAL INFORMATION:

; APPLICANT: WHYTE, DAVID

; APPLICANT: MANNING, GERARD

; APPLICANT: CAENEPEEL, SEAN

; TITLE OF INVENTION: NOVEL KINASES

; FILE REFERENCE: 034536-0321

; CURRENT APPLICATION NUMBER: US/10/618,941

; CURRENT FILING DATE: 2003-07-15

; PRIOR APPLICATION NUMBER: 60/395,632

; PRIOR FILING DATE: 2002-07-15

; NUMBER OF SEQ ID NOS: 143

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 55

; LENGTH: 4429

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-618-941-55

Alignment Scores:

| Pred. No.: | Score: | Percent Similarity: | Best Local Similarity: | Query Match: | DB: |
|------------|---------|---------------------|------------------------|---------------|-----|
| 0 | 2888.00 | 100.00% | 100.00% | 100.00% | 19 |
| Length: | 4429 | Matches: | 537 | Conservative: | 0 |
| | | Mismatches: | 0 | Indels: | 0 |
| | | Gaps: | 0 | | 0 |

US-10-631-958-10 (1-537) x US-10-618-941-55 (1-4429)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20

Db 92 ATGGGGCGCAGCGGGCGCGGAGCCGCTGCAATCCGTCTGTGGGTGAAGCAGCAGCGC 151

Qy 21 CyeAlaValSerLeuGluProAlaAArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40

Db 152 TGCGCCGTGAGCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGTGGCGAGCCCCGGGGCCC 211

Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCyserValProValSerGluIlelleAlaVal 60

Db 212 GGAGCCGGCGCCCCCGCGGGATGCTCTGTGCTGTATCTGAGATCATCGCGCTT 271

Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80

Db 272 GAGGAACACAGCGTTTACGCGGAAACATCAGGCAGTGGAAAATGGCAGAAAAATGGAAG 331

Qy 81 ProTyralaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100

Db 332 CCTTACCGCTTTTACAGTTCATCTGTGTAAAGAGACAGCAGCGCACCGCTGGAAAGTGGCG 391

Qy 101 GluValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120

Db 392 CAGGTGACTTTCGTGGTGTCCAGAGGACGCTGTGTACATTGTGTGGCTGCGAGACCCTGCGG 451

Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnPropHe 140

Db 452 GAGATGCTGGAGAGCTGACGTCAGACCAAAGCATTTACTGGTATTATTAACCCGTTTT 511

Qy 141 GlycylLysGlyGlnGlyLysArgIleTyrgluargLysValalaProLeupheThrLeu 160

Db 512 GGAGAAAAAGACAAGCAAGCGGATATATGAAGAAAGATGGCACCACCTGTTCACCTTA 571

| | | | |
|----|------|-----------------------------------------------------------------|------|
| Qy | 161 | AlaSerIleThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu | 180 |
| Db | 572 | GCCTCCATCACCTGACATCATCGTTACTGAACATGCTAATCAGGCAAGGAGACTCTG | 631 |
| Qy | 181 | TyrGluIleAsnIleAspIleValAspGlyIleValCysValGlyGlyAspGlyMetPhe | 200 |
| Db | 632 | TATGAGATTAAACATAGACAAATACGCGGCATCGTCTGTCTCGCGGAGATGGTATGTC | 691 |
| Qy | 201 | SerGluValLeuHisGlyLeuIleGlyValArgThrGlnArgSerAlaGlyValAspGlnAsn | 220 |
| Db | 692 | AGCGAGGTGCTGACCGTCTGATTGGGAGACGACGAGGAGCGCGGGGTGACACGAAC | 751 |
| Qy | 221 | HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer | 240 |
| Db | 752 | CACCCCGGGCTGTGCTGGTCCGAGTAGCTCCGATTGGAATCATTTCCGCGAGGGTCA | 811 |
| Qy | 241 | ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis | 260 |
| Db | 812 | ACGAGACTGCGTGTGTTACTCCACCGTGGGCACGACGACGACGAAACCTCGGCGCTGCAT | 871 |
| Qy | 261 | IleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSerThrLeu | 280 |
| Db | 872 | ATGCTGTTGGGAGCTCGCTGGCCATGGATGTCTCTCAGTCCACCACCAACGACACACTC | 931 |
| Qy | 281 | LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer | 300 |
| Db | 932 | CTTCGCTACTCCGTGTCCCTGTCTGGGTACGGCTCTCTACGGGACATCATCAAGGACAGT | 991 |
| Qy | 301 | GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu | 320 |
| Db | 992 | GAGAAGAAACGGTGGTGGTCTTGGCAGATACGACTTTTCAGGTTTTAAAGACCTTCCTC | 1051 |
| Qy | 321 | SerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer | 340 |
| Db | 1052 | TCCACCACTGCTATGAAGGAGCAGTGTCTCTCTCCCTGCACAAACACACGGTGGGATCT | 1111 |
| Qy | 341 | ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln | 360 |
| Db | 1112 | CCAAGGATAGGAAGCCCTGCGCGGAGGATGCTTTGTTTGACGGCAACCAAGCAGCAG | 1171 |
| Qy | 361 | LeuGluGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu | 380 |
| Db | 1172 | CTGAGGAGGAGCAGAAAGACACTGTATGTGTTTGGAACTCGCGAGGACGTGGAGCAG | 1231 |
| Qy | 381 | TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys | 400 |
| Db | 1232 | TGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCCAAAACATGCTCTGTGCTGT | 1291 |
| Qy | 401 | ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu | 420 |
| Db | 1292 | CGCGGAGCCCCAGGGGCTCTCCCGGGCTGCCACTTGGGAGACGGGTCTTTCTGACCTC | 1351 |
| Qy | 421 | IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn | 440 |
| Db | 1352 | ATCCTCATCCGAAAATGCTCCAGGTTCATTTTCTGAGATTCTCTCATCGGCACACCAAC | 1411 |
| Qy | 441 | GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe | 460 |
| Db | 1412 | CACGAGACACAGTTTGACTTTCATTTTTTTGGAAGTTTATCGCGTCAAGAAATTTCCAGTTT | 1471 |
| Qy | 461 | ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe | 480 |
| Db | 1472 | ACGTGGAAGCACATGGAGGATGAGGACGCGCTCAAGAGGGGGGGAAGAAGCGCTTT | 1531 |
| Qy | 481 | GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn | 500 |
| Db | 1532 | GGGCACATTTCAGCAGCCACCCCTCTCTGTGCTGCACCGTCTCTCAACAGCTCTCTGGAAC | 1591 |
| Qy | 501 | CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal | 520 |
| Db | 1592 | TGCGAGGGGAGGTCTGCACAGCCCTGCATCGAGTCCAGTCCAGTCCACTGCCAGCTGGTT | 1651 |

QY 521 ArgLeuPheAlaArgGlyIleGluGluAanProLysProAspSerHisSer 537
 Db 2764 CGACTCTTTGCACGAGGAATTGAAGAAATCCGAAGCCAGACTCACACAGC 2814

RESULT 8
 US-10-315-597A-1
 ; Sequence 1, Application US/10315597A
 ; Publication No. US20030162206A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sugiyura, Masaaki
 ; APPLICANT: Kono, Keita
 ; APPLICANT: Kohama, Takafumi
 ; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It
 ; FILE REFERENCE: 02658C1P/HG
 ; CURRENT APPLICATION NUMBER: US/10/315,597A
 ; PRIOR FILING DATE: 2002-12-10
 ; PRIOR APPLICATION NUMBER: JP 2000-178039
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 4463
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: 3371
 ; OTHER INFORMATION:
 US-10-315-597A-1

Alignment Scores:
 Pred. No.: 0 Length: 4463
 Score: 2880.00 Matches: 535
 Percent Similarity: 99.81% Conservative: 1
 Best Local Similarity: 99.63% Mismatches: 1
 Query Match: 99.72% Indels: 0
 DB: 16 Gaps: 0

US-10-631-958-10 (1-537) x US-10-315-597A-1 (1-4463)

QY 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
 Db 124 ATGGGGGCGACGGGGCGGGCGGAGCGCGTGCATTCGCTGTGGTGAAGCAGCGCGC 183

QY 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
 Db 184 TCGCGCGTGAGCGCTGGAGCGCGCGGGGCTCTGCTGGCTGGTGGCGAGCGCCGGGGCCC 243

QY 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 60
 Db 244 GGAGCGGGCGCGCGCGCGCGGATGCTGCTGTGCTGTATCTGAGATCATCGCCGTT 303

QY 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
 Db 304 GAGGAACAGACGTTACCGGGAACATCAAGCAGTGGAAATGGCAAAATGGAAAG 363

QY 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100
 Db 364 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGACGACGCGCACCGCTGGAGTGGGG 423

QY 101 GlnValThrPheTrpCysProGluGluGlnLysHisLeuTrpLeuGlnThrLeuArg 120
 Db 424 CAGGTGACTTTCTGTGTGTCCAGAGGAGCAGCTGTCTGCTGTGCTGTATCTGAGATCATCGCCGTT 483

QY 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
 Db 484 GAGATGCTGGGAAGCTGACGTCCAGACCAAAAGCATTTTACTGGTATTATCAACCCGTTT 543

QY 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
 Db 544 GGAGGAAAGGACAGGCGGCGGATATGAAAGAAAGTGGCACCACTGTTCACCTTA 603

QY 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAenGlnAlaLysGluThrLeu 180
 Db 604 GCCTCCATCACCACCTGACATCATCGTTACTGAACATGCTAATCAGGCCAAGGAGACTCTTG 663

QY 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
 Db 664 TATGAGATTACATAGACAAATACACCGCATCGTCTGTTCGGCGGAGATGGTATGTTTC 723

QY 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAen 220
 Db 724 AGCAGGCTGCTGCACGGTCTGATTGGGAGGACGACGAGAGCGCGCGGTCCACAGAAC 783

QY 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
 Db 784 CACCCCGGGGTGTGCTGCTCCCAAGTAGCTCCCGATTGGAAATCATTTCCCGCAGGTCCA 843

QY 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
 Db 844 ACGGACTGCGTGTGTACTCCACCGTGGGACGACGAGCGAGAAACCTTCGGCGGTGCAT 903

QY 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
 Db 904 ATCGTTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAGCACATC 963

QY 281 LeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIleLysAspSer 300
 Db 964 CTTGCTACTCCGTGCTCCGCTACGGCTTCTACGGGACATCATCAAGACAGT 1023

QY 301 GluLysLysArgTrpLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
 Db 1024 GAGAAGAAACGGTGTGGTCTTGCCAGATACGACTTTTCAAGTTTAAAGACCTTCCTC 1083

QY 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
 Db 1084 TCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCGACACACACACGCTGGGATCT 1143

QY 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
 Db 1144 CCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGAGGCAAGCAAGCAGCAG 1203

QY 361 LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
 Db 1204 CTGGAGGAGGAGCAGAAAGACCTGTATGTTTGAAGCTGCGGAGAGCTGGAGGAG 1263

QY 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
 Db 1264 TGGCAAGTCGTCGTGGGAAGTTCTGGCCATCAATGCCACAAACATGTCCTGTCTGT 1323

QY 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
 Db 1324 CGCCGAGCCCGCAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC 1383

QY 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAen 440
 Db 1384 ATCTCATCCGAAATGCTCCAAAGTTCAATTTTCTGAGATTTCTCATCAGGCACCAAC 1443

QY 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
 Db 1444 CAGCAGGACCACTGAGTTCACCTTCTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1503

QY 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
 Db 1504 ACGTGCAAGCAGCAGGAGGATGAGACAGCAGCTCAAGGAGGGGGGGAAGAGCGCTTT 1563

QY 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAen 500
 Db 1564 GGGCACATTTCCAGCAGCACCCCTCTCTGCTGCTGACCGTCTCAACAGCTCTCTGGAAC 1623

QY 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
 Db 1624 TGGCATGGGAGGTCCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCAGCTCCAGCTGTT 1683

QY 521 ArgLeuPheAlaArgGlyIleGluGluAanProLysProAspSerHisSer 537

Db 1684 CGACTCTTTGACGAGGAATTAAGAGAAATCCGAAGCCAGACTCACACAGC 1734
RESULT 9
US-09-784-810A-5
; Sequence 5, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-784-810A-5
Alignment Scores:
Pred. No.: 8,37e-303 Length: 1840
Score: 2649.50 Matches: 495
Percent Similarity: 97.83% Conservative: 1
Best Local Similarity: 97.63% Mismatches: 0
Query Match: 91.74% Indels: 11
DB: 9 Gaps: 1
US-10-631-958-10 (1-537) x US-09-784-810A-5 (1-1840)
Qy 42 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaValGlu 61
Db 1 GCCGGCGCCCGCGCGGATGCCTCTGTGCTGTATCTGAGATCATCGCGTTGAG 60
Qy 62 GluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro 81
Db 61 GAAACAGACGTTCAACGGGAAACATCAAGCGCATGTGGAAATGGCAAAATGGAAAGCCT 120
Qy 82 TyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAlaGln 101
Db 121 TACGCTTTTACAGTTCTACTGTGTAAAGAGACACGCGCACCCCTGGAAAGTGGCGGAG 180
Qy 102 ValThrPheTrpCysProGluGluGlnLeuLysHisLeuTrpLeuGlnThrLeuArgGlu 121
Db 181 GTGACTTTCTGTTGCCAGAGGAGCAGCTGTGTCTACCTTGTGGCTGCAGACCTCGCGGAG 240
Qy 122 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 141
Db 241 ATGCTGGAGAAAGCTGACGCTCCAGACCAAGCATTTTACTGGTATTTATCAACCCGTTTGA 300
Qy 142 GlyLysGlyGlnGlyLysArgIleTrpGluArgLysValAlaProLeuPheThrLeuAla 161
Db 301 GGAAGAGACAAGCGACCGGATATATGAAAGAAAGTGGCACCCACTGTTCCACTTTAGCC 360
Qy 162 SerIleThrAspIleIle-----ValThr 170
Db 361 TCCATCACCCTGACATCATCGGTAAACAAATTCCTATGTTAACTATGTAGAGTAATTACT 420
Qy 171 GluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTrpAspGly 190
Db 421 GAAATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCGC 480
Qy 191 IleValCysValGlyLysGlyMetPheSerGluValLeuHisGlyLysGlyValArg 210
Db 481 ATCGTCTGTGCGCGGAGATGGTATGTTTACGAGAGGTGCTGCAGCGTCTGATTTGGAGG 540
Qy 211 ThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSer 230

Db 541 ACGCAGAGGAGCGCGCGGTCGACACAGAACACCCCGGGCTGTGCTGGTCCCCAGTAGC 600
Qy 231 LeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTrpSerThrValGly 250
Db 601 CTCGGATTGGAATCATTTCCCGAGGGTCAACGAGCTGCGTGTGTACTCCACCGTGGC 660
Qy 251 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAsp 270
Db 561 ACCAGCAGCAGAGAACTCGGCGCTGCATATCGTTGTGGGACTCGCTGGCCATGGAT 720
Qy 271 ValSerSerValHisAsnSerThrLeuLeuArgTrpSerValSerLeuGlyTyr 290
Db 721 GTGTCTCTAGTCCACCAACACAGCACACTCTTCCTCGCTACTCCGTGCTCCCTGCTGGGCTAC 780
Qy 291 GlyPheTrpGlyAspIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArg 310
Db 781 GGCTTCTACCGGGACATCAAGGACAGTGAAGAAACGGTGGTGGGTCTTTGCCAGA 840
Qy 311 TyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTrpGluGlyThrValSer 330
Db 841 TACGACTTTTACAGTTTAAAGACTTCTCTCCACCACCTGCTATGAAGGACAGTGTCC 900
Qy 331 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 350
Db 901 TTCTCTCCCTGCACAAACACACGCTGGGATCTCCAAGGATAGGAAGCCCTCGCGGCAGGA 960
Qy 351 CysPheValCysArgGlnSerLysGlnLeuGluGluGlnLysLysAlaLeuTyr 370
Db 961 TGCTTTGTTTGCAGGGCAAGCAAGCAGCAGCTGGAGAGGAGGACAGAAAGCACTGTAT 1020
Qy 371 GlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAla 390
Db 1021 GGTTTGGAAGCTCGCGAGGACGTGGAGAGTGGCAAGTCTCTGTGGGAAGTTTCTGCC 1080
Qy 391 IleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAla 410
Db 1081 ATCAATGCCACAAACATGCTCTGTGCTGTGCGCGAGCCCGAGGGGCTCTCCCGCGCT 1140
Qy 411 AlaHisLeuGlyAspGlySerSerAspLeuIleuIleArgLysCysSerArgPheAsn 430
Db 1141 GCCCAGCTGGGAGACGGGTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTCAAT 1200
Qy 431 PheLeuArgPheLeuIleArgHisThrAsnGlnAspGlnPheAspPheThrPheVal 450
Db 1201 TTTCTGAGATTTCTCATCAGGCACACCAACACAGCAGACCCAGTTTGACTTCTTTGTT 1260
Qy 451 GluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSer 470
Db 1261 GAAGTTTATCGCGTCAAGAAATTCAGTTTACGTTCGAAGCACATGGAGGATGAGGACAGC 1320
Qy 471 AspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCys 490
Db 1321 GACCTCAAGAGGGGGGAAAGAGCGCTTTGGGCACATTTGACAGACCCACCTCTCTGC 1380
Qy 491 CysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAla 510
Db 1381 TGCTGACCGCTCTCCAACAGCTCTTGGAACTGGCAGCGGGAGGTCTCTGCACAGCCCTGCC 1440
Qy 511 IleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsn 530
Db 1441 ATCAGGTCAGGGTCCACTGCCAGCTGCTTTCGACTCTTTCGACGAGGAATTAAGAGAAT 1500
Qy 531 ProLysProAspSerHisSer 537
Db 1501 CCGAAGCCAGACTCACACAGC 1521
RESULT 10
US-10-876-281-5
; Sequence 5, Application US/10876281
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA

| | | |
|----|--------------------------------------------------------------------------|-----|
| ; | TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING | |
| ; | TITLE OF INVENTION: SAME | |
| ; | FILE REFERENCE: 10716-08 | |
| ; | CURRENT APPLICATION NUMBER: US/10/876,281 | |
| ; | CURRENT FILING DATE: 2004-06-24 | |
| ; | PRIOR APPLICATION NUMBER: US/09/784,810 | |
| ; | PRIOR FILING DATE: 2001-02-14 | |
| ; | PRIOR APPLICATION NUMBER: 60/182,360 | |
| ; | PRIOR FILING DATE: 2000-02-14 | |
| ; | PRIOR APPLICATION NUMBER: 60/191,261 | |
| ; | PRIOR FILING DATE: 2000-03-22 | |
| ; | NUMBER OF SEQ ID NOS: 29 | |
| ; | SOFTWARE: PatentIn Ver. 2.1 | |
| ; | SEQ ID NO 5 | |
| ; | LENGTH: 1840 | |
| ; | TYPE: DNA | |
| ; | ORGANISM: Homo sapiens | |
| ; | US-10-876-281-5 | |
| ; | | |
| ; | Alignment Scores: | |
| ; | Pred. No.: 8,37e-303 Length: 1840 | |
| ; | Score: 2649.50 Matches: 495 | |
| ; | Percent Similarity: 97.83% Conservative: 1 | |
| ; | Best Local Similarity: 97.63% Mismatches: 0 | |
| ; | Query Match: 91.74% Indels: 11 | |
| ; | DB: 22 Gaps: 1 | |
| ; | | |
| ; | US-10-631-958-10 (1-537) x US-10-876-281-5 (1-1840) | |
| Qy | 42 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaValGlu | 61 |
| Db | 1 GCGGGCGCCCGCGCGCGATGCGCTGCTGTGCGCTGTATCTGAGATCATCGCGGTTGAG | 60 |
| Qy | 62 GluThrAspValHisGlyLyHisGlnGlySerGlyLySstTpGlnLySstMetGluLySPro | 81 |
| Db | 61 GAAACAGACGTTACGGGAAACATCCAGCGAGTGGAAATGGCAGAAATGGAAAGCCCT | 120 |
| Qy | 82 TyrAlaPheThrValHisCysVallysArgAlaArgHisArgTstTpLysTrpAlaGln | 101 |
| Db | 121 TAGCGCTTTTACAGTTTCACTGTGTAAAGAGACGACGGCACCGCTGGAAGTGGCGCGAG | 180 |
| Qy | 102 ValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu | 121 |
| Db | 181 GTGACTTTCGTGTGTCCAGAGAGCAGCTGTGTCTACCTTGTGGCTGCAGACCCCTGCGGGAG | 240 |
| Qy | 122 MetLeuGluLySLeuThrSerArgProLyHisLeuLeuValPheIleAenProPheGly | 141 |
| Db | 241 ATGCTGGAGAGCTCAGCTGCAGACCAAGACATTTACTGGTATTTATCAACCCGTTTGA | 300 |
| Qy | 142 GlyLySgLyGlnGlyLeArgIleTytrGluArgLySValAlaProLeuPheThrLeuAla | 161 |
| Db | 301 GGAAGAGGACAGGACAGCGGATATATGAAGAAAAGTGGCACCACTGTTCACTTAGCC | 360 |
| Qy | 162 SerIleThrThrAspIleIle-----ValThr | 170 |
| Db | 361 TCCATCACCACCTGCATCATCGTTAAACAAATTTCTATGTTAACTATGTAGAAGTAATTA | 420 |
| Qy | 171 GluHisAlaAsnGlnAlaLySgLuThrLeuTytrGluIleAenIleAspLySstTrpAspGly | 190 |
| Db | 421 GAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGC | 480 |
| Qy | 191 IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArg | 210 |
| Db | 481 ATGCTCTGTGTGGCGGAGATGGTATGTTTCAGCAGAGTGTCTGCACGGTCTGATTTGGAG | 540 |
| Qy | 211 ThrGlnArgSerAlaGlyValAspGlnAenHisProArgAlaValLeuValProSerSer | 230 |
| Db | 541 ACGCAGAGGAGCGCGGGTGCACAGAAACACACCCCGGGCTGTGTGTGCTGTGCTGTGAG | 600 |
| Qy | 231 LeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTrpSerThrValGly | 250 |
| Db | 601 CTCGGATTGGAAATCATTTCCCGCGGGGTCAACGGAGCTGCGTGTGTATTCTCCACCGTGG | 660 |

```
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-1

Alignment Scores:
Pred. No.: 1.27e-183 Length: 979
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 56.80% Indels: 11
DB: 10 Gaps: 1

US-10-631-958-10 (1-537) x US-09-969-896-1 (1-979)

Qy 130 ProlysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 149
Db 2 CCAAGCATTTTACTGGTATTATCAACCGTTTGAGGAAAGGACAAAGGCGGATGATA 61

Qy 150 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle--- 168
Db 62 TATGAAAGAAAGTGGCACCACTGTTACCTTAGCCTCCATCACCCTGACATCATCGGT 121

Qy 169 -----ValThrGluHisAlaAsnGlnAlaLysGlu 178
Db 122 AACAAATTTCTATGTTAACTATGTAGAAGTAATTACTGAACATGCTAATCAGGCCAAGGAG 181

Qy 179 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 198
Db 182 ACTCTGATGAGATTACATAGACAAATACGCGCATCGTCTGTGTGCGGAGATGGT 241

Qy 199 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 218
Db 242 ATGTTACGAGGAGTGTGCACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGGTCCAC 301

Qy 219 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 238
Db 302 CAGAACCAACCCCGGGCTGCTGGTCCCAAGTAGTACCGCTCCGATTTGAATCATTTCCCGCA 361

Qy 239 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 258
Db 362 GGGTCACCGACTGCGTGTGTACTCCACCGTGGCCACGACGACGAGAACCTCGCG 421

Qy 259 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 278
Db 422 CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACACAGC 481

Qy 279 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleIys 298
Db 482 ACACCTCTCTACTACTCGCTGCTGCTGGGTACGGCTTCTACGGGACATCATCAAG 541

Qy 299 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 318
Db 542 GACAGTGAGAGAAACGGTGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 601

Qy 319 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 338
Db 602 TTCCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGTCACACACACACGGTG 661

Qy 339 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIys 358
Db 662 GGATCTCCAGGGATAGGAAGCCCTGCGGCGGAGATGCTTTGTTTTCAGGCAAGCAAG 721

Qy 359 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 378
Db 722 CAGCAGCTGGAGGAGGACAGAGAACCATCTGTATGTTTGGAGCTGCGGAGGACGTG 781

Qy 379 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaIleThrAsnMetSerCys 398
```

```
Db 782 GAGGAGTGGCAAGTCGTCTCTGGGAAGTTTCTGCCATCAATGCCAACAATGCTCTGT 841
Qy 399 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 418
Db 842 GCTTGTGCGCGAGCCCAAGGGGCTCTCCCGGCTGCCCATCTGGGAGACGGGTCTTCT 901
Qy 419 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 438
Db 902 GACCTCATCTCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCAC 961

Qy 439 ThrAsnGlnGlnAspGln 444
Db 962 ACCAACCCAGCAGGACCAG 979

RESULT 12
US-10-631-958-1
; Sequence 1, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631.958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-1

Alignment Scores:
Pred. No.: 1.27e-183 Length: 979
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 56.80% Indels: 11
DB: 19 Gaps: 1

US-10-631-958-10 (1-537) x US-10-631-958-1 (1-979)

Qy 130 ProlysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 149
Db 2 CCAAGCATTTTACTGGTATTATCAACCGTTTGAGGAAAGGACAAAGGCGGATGATA 61

Qy 150 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle--- 168
Db 62 TATGAAAGAAAGTGGCACCACTGTTACCTTAGCCTCCATCACCCTGACATCATCGGT 121

Qy 169 -----ValThrGluHisAlaAsnGlnAlaLysGlu 178
Db 122 AACAAATTTCTATGTTAACTATGTAGAAGTAATTACTGAACATGCTAATCAGGCCAAGGAG 181

Qy 179 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 198
Db 182 ACTCTGATGAGATTACATAGACAAATACGCGCATCGTCTGTGTGCGGAGATGGT 241

Qy 199 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 218
Db 242 ATGTTACGAGGAGTGTGCACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGGTCCAC 301

Qy 219 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 238
Db 302 CAGAACCAACCCCGGGCTGCTGGTCCCAAGTAGTACCGCTCCGATTTGAATCATTTCCCGCA 361

Qy 239 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 258
Db 362 GGGTCACCGACTGCGTGTGTACTCCACCGTGGCCACGACGACGAGAACCTCGCG 421

Qy 259 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 278
Db 422 CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACACAGC 481

Qy 279 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleIys 298
Db 482 ACACCTCTCTACTACTCGCTGCTGCTGGGTACGGCTTCTACGGGACATCATCAAG 541

Qy 299 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 318
Db 542 GACAGTGAGAGAAACGGTGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 601

Qy 319 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 338
Db 602 TTCCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGTCACACACACACGGTG 661

Qy 339 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIys 358
Db 662 GGATCTCCAGGGATAGGAAGCCCTGCGGCGGAGATGCTTTGTTTTCAGGCAAGCAAG 721

Qy 359 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 378
Db 722 CAGCAGCTGGAGGAGGACAGAGAACCATCTGTATGTTTGGAGCTGCGGAGGACGTG 781

Qy 379 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaIleThrAsnMetSerCys 398
```

```
QY 239 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 258
DB 362 GGGTCAACGGACTGGTGTGTACTCCACCGTGGGCACCGACGACGAGAACCTCGCGC 421
QY 259 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 278
DB 422 CTGCATATCGTGTGGGACTCGTGGCCATGGATGTCTCAGTCCACCAACACAGC 481
QY 279 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuys 298
DB 482 ACACCTCCTTCGTACTCCGCTGCTGTGGCTACCGCTTCTACGGGACATCATCAAG 541
QY 299 AspSerGluLysLysArgTyrPheLysLysAlaArgTyrAspPheSerGlyLeuLysThr 318
DB 542 GACAGTGAAGAAACGGTGGTGTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 601
QY 319 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 338
DB 602 TTCCTCTCCACCATGCTATGAGAGGACAGTGTCTCTCTCCCTGCAACACACAGGTG 661
QY 339 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 358
DB 662 GGATCTCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGAGGCAAGCAAG 721
QY 359 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 378
DB 722 CAGCAGCTGGAGGAGGAGCAGAAAGCACTGTATGTTTGGAGCTGCGGAGGACGTG 781
QY 379 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 398
DB 782 GAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACCAATGTCTCTGT 841
QY 399 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 418
DB 842 GCTTGTCCCGAGAGCCCGAGGGCTCTCCCGGGCTGCCACTTGGGAGCGGGTCTTCT 901
QY 419 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 438
DB 902 GACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 961
QY 439 ThrAsnGlnGlnAspGln 444
DB 962 ACCAACGACGAGGACCAG 979
RESULT 13
US-10-115-635-247
; Sequence 247, Application US/10115635
; Publication No. US20040137434A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Weinman, Tom
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 797CON
; CURRENT APPLICATION NUMBER: US/10/115,635
; PRIORITY FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 09/714,936
; NUMBER OF SEQ ID NOS: 362
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 247
; LENGTH: 817
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (90)..(662)
US-10-115-635-247
Alignment Scores:
Pred. No.: 1,29e-111 Length: 817
Score: 1032.00 Matches: 190
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.73% Indels: 0
DB: 19 Gaps: 0
US-10-631-958-10 (1-537) x US-10-115-635-247 (1-817)
QY 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTyrValLysGlnGlnArg 20
DB 90 ATGGGGGGGACCGGGGGGGGGAGCCGCTGTCGATCCGTCGCTGGGTGAAGCAGCAGCGC 149
QY 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTrpArgSerProGlyPro 40
DB 150 TGGCCGCTGAGCGCTGGAGCCGCGCGGCTCTGCTGGCTGTCTGTGATCTGAGATCATCGCCGTT 269
QY 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 60
DB 210 GGAGCCGCGCGCCCGCGCGGATGCCTGCTCTGTGCTGTATCTGAGATCATCGCCGTT 269
QY 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTyrGlnLysMetGluLys 80
DB 270 GAGGAACAGACGTTACGGGAAACATCAAGGCAGTGGAAATGGCAAAATGGAAAG 329
QY 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTyrLysTrpAla 100
DB 330 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGACGCGCTGGAAGTGGGCG 389
QY 101 GlnValThrPheTyrCysProGluGlnLeuCysHisLeuTyrLeuGlnThrLeuArg 120
DB 390 CAGGTGATTTTCTCGTGTCCAGAGGAGCAGCTGTGTCTACCTTGTGGCTGCAGACCTCGCG 449
QY 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
DB 450 GAGATGCTGGAGAAGCTGACGTCCAGACCAAGCATTTTACTGGTATTTTATCAACCCGTT 509
QY 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
DB 510 GGAGAAAGGACAAAGCAAGCGGATATATGAAAGAAAGTGGCACCCTGTTCACCTTA 569
QY 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
DB 570 GCCTCCATCACCCTGACATCATCTTACTGAACATGCTATACAGGCCAAGGAGACTCTG 629
QY 181 TyrGluIleAsnIleAspLysTyrAspGly 190
DB 630 TATGAGATTAAATAGACAAATACGACGG 659
RESULT 14
US-09-969-896-8
; Sequence 8, Application US/09969896
; Publication No. US2003012553A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIORITY FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-8

Alignment Scores:
Pred. No.: 8.49e-93 Length: 550
Score: 871.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.16% Indels: 0
DB: 10 Gaps: 0

US-10-631-958-10 (1-537) x US-09-969-896-8 (1-550)

```
Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
Db 76 ATGGGGCGACGGGGCGGGCGGAGCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC 135
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 136 TGCCTCGTGAAGCTTGGAGCCCGCGGGCTCTGCTGCGCTGTGGTGAAGCAGCAGCGC 195
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluLeuIleAlaVal 60
Db 196 GGAGCCGGCGCCCGCGGGCGGATGCTGCTGTGCTGTATCTGAGATCATCGCCGTT 255
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 256 GAGGAAACAGACGTTCAAGGCAACATCAAGGCAGTGGAAATGGCAGAAATGAAAAAG 315
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100
Db 316 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGCCCGCTGGAAGTGGCG 375
Qy 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db 376 CAGGTGACTTTCTGGTGTCCAGAGGACGAGTGTGTCACCTTGTGGCTGCAGACCTGCGG 435
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 436 GAGATGCTGGAGAAAGCTGACGTCAGACCAAGCATTTACTGTATTATCAACCCGTTT 495
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPhe 158
Db 496 GGAGGAAAAGGACAAAGCGGATATATGAAAAGAAAAGTGGCACCCTGTTT 549
```

RESULT 15

US-10-631-958-8
; Sequence 8, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631.958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-8

Alignment Scores:
Pred. No.: 8.49e-93 Length: 550
Score: 871.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.16% Indels: 0
DB: 19 Gaps: 0

US-10-631-958-10 (1-537) x US-10-631-958-8 (1-550)

```
Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
Db 76 ATGGGGCGACGGGGCGGGCGGAGCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC 135
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 136 TGCCTCGTGAAGCTTGGAGCCCGCGGGCTCTGCTGCGCTGTGGTGAAGCAGCAGCGC 195
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluLeuIleAlaVal 60
Db 196 GGAGCCGGCGCCCGCGGGCGGATGCTGCTGTGCTGTATCTGAGATCATCGCCGTT 255
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 256 GAGGAAACAGACGTTCAAGGCAACATCAAGGCAGTGGAAATGGCAGAAATGAAAAAG 315
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100
Db 316 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGCCCGCTGGAAGTGGCG 375
Qy 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db 376 CAGGTGACTTTCTGGTGTCCAGAGGACGAGTGTGTCACCTTGTGGCTGCAGACCTGCGG 435
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 436 GAGATGCTGGAGAAAGCTGACGTCAGACCAAGCATTTACTGTATTATCAACCCGTTT 495
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPhe 158
Db 496 GGAGGAAAAGGACAAAGCGGATATATGAAAAGAAAAGTGGCACCCTGTTT 549
```

Search completed: September 6, 2005, 10:48:14
Job time : 1559.64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 5, 2005, 19:28:44 ; Search time 303.735 Seconds
(without alignments)
2892.921 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 MGATGAEPQLSVLWVKQR.....QLVRLFARGIENKPDPSHS 537

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10631958/runat_02092005_165811_4432/app_query.fasta_1.1941
-DB=Issued Patents NA -OPMT=fastap -SUPFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CGN_1_1_397 @runat_02092005_165811_4432 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/FACTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 2881 | 99.8 | 4432 | 4 | US-09-774-528-148 |
| 2 | 801.5 | 27.8 | 2064 | 4 | US-09-270-767-14306 |
| 3 | 585.5 | 20.3 | 1084 | 4 | US-09-270-767-15155 |
| 4 | 302.5 | 10.5 | 901 | 4 | US-09-270-767-30448 |
| 5 | 278 | 9.6 | 1857 | 4 | US-09-970-516-3 |
| 6 | 278 | 9.6 | 2380 | 4 | US-09-817-676A-13 |
| 7 | 269 | 9.3 | 2898 | 4 | US-09-817-676A-11 |
| 8 | 266.5 | 9.2 | 1205 | 4 | US-09-959-897-1 |
| 9 | 264.5 | 9.2 | 1155 | 4 | US-09-970-516-1 |
| 10 | 263.5 | 9.1 | 1783 | 4 | US-09-949-016-1155 |
| 11 | 257.5 | 8.9 | 1533 | 4 | US-09-205-258-90 |
| 12 | 247.5 | 8.6 | 1149 | 4 | US-09-970-516-5 |

| | | | | | |
|----|-------|-----|---------|---|---------------------|
| 13 | 228 | 7.9 | 1611 | 4 | US-09-248-796A-1756 |
| 14 | 210.5 | 7.3 | 1875 | 4 | US-09-614-221A-399 |
| 15 | 203.5 | 7.0 | 1764 | 4 | US-09-949-016-3940 |
| 16 | 199.5 | 6.9 | 2462 | 4 | US-09-620-312D-796 |
| 17 | 184 | 6.4 | 1012 | 4 | US-09-270-767-12677 |
| 18 | 153 | 5.3 | 942 | 4 | US-09-107-532A-3500 |
| 19 | 153 | 5.3 | 1239 | 4 | US-09-603-208A-263 |
| 20 | 144 | 5.0 | 959 | 4 | US-09-270-767-6490 |
| 21 | 144 | 5.0 | 959 | 4 | US-09-270-767-21772 |
| 22 | 127 | 4.4 | 7195 | 4 | US-09-949-016-12897 |
| 23 | 127 | 4.4 | 7198 | 4 | US-09-949-016-15682 |
| 24 | 125 | 4.3 | 485 | 4 | US-09-270-767-31476 |
| 25 | 123.5 | 4.3 | 915 | 4 | US-09-134-000C-1961 |
| 26 | 119.5 | 4.1 | 3089 | 4 | US-09-634-238A-188 |
| 27 | 118 | 4.1 | 912 | 4 | US-09-107-532A-1479 |
| 28 | 118 | 4.1 | 4411529 | 3 | US-09-103-840A-1 |
| 29 | 116.5 | 4.0 | 885 | 4 | US-09-583-110-1079 |
| 30 | 116.5 | 4.0 | 948 | 4 | US-09-107-433-518 |
| 31 | 113.5 | 3.9 | 1035 | 4 | US-09-134-000C-408 |
| 32 | 113 | 3.9 | 2925 | 4 | US-09-883-134-6 |
| 33 | 112.5 | 3.9 | 2612 | 4 | US-09-949-016-4164 |
| 34 | 111.5 | 3.9 | 900 | 3 | US-08-961-527-332 |
| 35 | 111.5 | 3.9 | 93398 | 4 | US-09-949-016-14167 |
| 36 | 110 | 3.8 | 975 | 3 | US-09-134-001C-1937 |
| 37 | 110 | 3.8 | 3761 | 4 | US-09-710-279-3656 |
| 38 | 110 | 3.8 | 10717 | 4 | US-09-902-540-991 |
| 39 | 109.5 | 3.8 | 36941 | 4 | US-08-311-731A-130 |
| 40 | 109 | 3.8 | 498 | 4 | US-09-893-737-59 |
| 41 | 107.5 | 3.7 | 978 | 4 | US-09-107-532A-3132 |
| 42 | 107 | 3.7 | 191 | 4 | US-09-270-767-28485 |
| 43 | 106 | 3.7 | 4411529 | 3 | US-09-103-840A-1 |
| 44 | 105.5 | 3.7 | 2469 | 4 | US-09-252-991A-5908 |
| 45 | 105.5 | 3.7 | 3135 | 4 | US-09-252-991A-5922 |

ALIGNMENTS

RESULT 1
US-09-774-528-148
; Sequence 148, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Duanrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/774, 528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 148
; LENGTH: 4432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129) ..(2817)
US-09-774-528-148
Alignment Scores:

```
Pred. No.: 1.18e-317 Length: 4432
Score: 2881.00 Matches: 536
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 99.76% Indels: 0
DB: 4 Gaps: 0

US-10-631-958-10 (1-537) x US-09-774-528-148 (1-4432)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20
Db |||||
Qy 1204 ATGGGGGCGAGCGGGCGCGGAGCCGCTGCAATCCGTCGTGGGTGAAGCAGCAGCGC 1263
Db |||||
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db |||||
Qy 1264 TGGCCGCTGAGCTGGAGCCCGCGGGGCTCTGTCGCTGTGTCGGAGCCCGGGGCC 1323
Db |||||
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db |||||
Qy 1324 GGAGCCGGCGCCCGCGCGGATGCTGCTCTGTCGTATCTGAGATCATGCCGTT 1383
Db |||||
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db |||||
Qy 1384 GAGGAAACAGACGTTCCAGCGGAAACATCAAGGCAGTGGAAATGGCAGAAATGGAAG 1443
Db |||||
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100
Db |||||
Qy 1444 CCTTACGCTTTACAGTTCACTGTGTAAAGAGACGACGCGCACCGTGGAAAGTGGCG 1503
Db |||||
Qy 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db |||||
Qy 1504 CAGGTGACTTTCTGGTGTCCAGAGAGCAGCTGTGTCACTTGTGGCTGCAGACCTGCGG 1563
Db |||||
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db |||||
Qy 1564 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGGTATTATCAACCCGTTT 1623
Db |||||
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db |||||
Qy 1624 GGAGGAAAGACCAAGCAAGCGGATATGAAAGAAAGTGGCACCACTGTTTCAACCTTA 1683
Db |||||
Qy 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db |||||
Qy 1684 GCCTCCATCACTACATCATCGTTACTGAAATGCTAATCAGGCAAGGAGACTCTG 1743
Db |||||
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db |||||
Qy 1744 TATGAGATTAAATAGCAAAATACAGCGCATCGTCTGTGTCGGCGGAGATGGTATGTT 1803
Db |||||
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db |||||
Qy 1804 AGCGAGGTGCTGCAGGCTCTGATTGGGAGGACGACAGAGCGCGGGGTCGACCAAGAC 1863
Db |||||
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db |||||
Qy 1864 CACCCCGGGCTGTGCTGCTCCAGTAGCTCCGGATTGGAATCATTTCCCGCAGGGTCA 1923
Db |||||
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db |||||
Qy 1924 ACGGACTCGGTGTATTCTCCACCGTGGGACCGACGAGCAAACTCGCGCGCTGCAT 1983
Db |||||
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
Db |||||
Qy 1984 ATCGTGTGGGACTCGCTGCCATGGATGTGCTCCTCAGTCCACCAACAGCACACTC 2043
Db |||||
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
Db |||||
Qy 2044 CTTGCTACTCCGTCTCCCTGCGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT 2103
Db |||||
Qy 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
Db |||||
Qy 2104 GAGAAGAAACGGTGGTGGGTTCTGGCCAGATACACTTTTTCAGGTTTAAAGACCTTCTC 2163
Db |||||

321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db |||||
2164 TCCACCACCTGCTATGGAGGACAGTGTCTCTCCTCCACCAACACACGGTGGGATCT 2223
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db |||||
2224 CCAAGGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTTGCAGGCAACAGCAGCAG 2283
Qy 361 LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db |||||
2284 CTGAGGAGGAGGACGAAGAAGCAGTGTATGGTTTGGAACTCGCGAGGACGTGGAGGAG 2343
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db |||||
2344 TGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCACAAACATGTCTGTGCTTGT 2403
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db |||||
2404 CGCCGGAGCCCAAGGGGCTCTCCCGGCTGCCACCTTGGGAGACGGGTCTTCTGACCTC 2463
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db |||||
2464 ATCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAAC 2523
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db |||||
2524 CAGCAGGACCACTTGAATTTCACTTTTGTGTAAGTTTATCCGCTCAAGAAATCCAGTTT 2583
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
Db |||||
2584 ACCTCGAAGCACATGGAGGATGAGCAGCAGCTCAAGGAGGGGGGAAAGCGCTTT 2643
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
Db |||||
2644 GGGCACATTTGACGACGCCACCTCTCTGCTGTGACCCGCTTCCACACACTCTCTGGAAC 2703
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db |||||
2704 TGGCAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCCAGTCCACTGCCAGTGGTT 2763
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
Db |||||
2764 CGACTCTTTCCAGGAGGATTTGAAGAGATCCGAAGCCAGACTCACACAGC 2814

RESULT 2
US-09-270-767-14306
; Sequence 14306, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14306
; LENGTH: 2064
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14306

Alignment Scores:
Pred. No.: 1.02e-80 Length: 2064
Score: 801.50 Matches: 197
Percent Similarity: 46.74% Conservative: 97
Best Local Similarity: 31.32% Mismatches: 206
Query Match: 27.75% Indels: 129
DB: 4 Gaps: 16

US-10-631-958-10 (1-537) x US-09-270-767-14306 (1-2064)

Qy 11 GlnSerValLeuTrpValLysGlnGlnArgCysAlaValSerLeuGlu-----ProAla 28
```



```
Pred. No.: 1,56e-56 Length: 1084
Score: 585.50 Matches: 125
Percent Similarity: 51.14% Conservative: 55
Best Local Similarity: 35.51% Mismatches: 103
Query Match: 20.27% Indels: 69
DB: 4 Gaps: 6

US-10-631-958-10 (1-537) x US-09-270-767-15155 (1-1084)

Qy 129 ArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyArg 148
Dy 1044 CGAGTGGCCGCCCTTTGGTCTTTATATAACCCCTATGGAGGTGCGAAGCGGGAGCTCAG 985
Qy 149 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIle 168
Dy ACCTATGAGCCCATGTGAGACCTATTTTCAGCTTGGCGCGTAGAGCCACCGTGTATC 925
Qy 169 ValThrGluHisAlaAsnGlnAlaGlySerLeuTyrGluIleAsnIleAspLysTyr 188
Dy ACCACTCAGAGGGCAACCAAGTGAAGGACATCTTCTGAGCCATGATCTGGGAGTATAC 865
Qy 189 AspGlyIleValCysValGlyCysGlyMetPheSerGluValLeuHisGlyLeuIle 208
Dy GATCGGTTTCTGTGTCGAGGCGATGGCACCGTACGAGAGGTTCATCAACGGACTGATA 805
Qy 209 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 228
Dy TTCGCTCAATGCGAGAGTGGAGTGGAGTGGAGACAGCGGCCACCATCATCTCCAGACCG 745
Qy 229 SerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTyrSerThr 248
Dy ---GCTCGCAGTGGGTGTATTCCTCCGCTGGCAGCACCGACACCATTTGCGTATAGTATG 688
Qy 249 ValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla 268
Dy CACGCGACGGCGATGTGAGGACACGCGCTATCCATGTATTCTGGCGCACCATCGGGGA 628
Qy 269 MetAspValSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeu 288
Dy TTGATGTGTGATGTGAGCAATGGCCAGTCCCTGCTCAGATCTGTGCGCAGTGTCTGTG 568
Qy 289 GlyTyrGlyPheTyrGlyAspIleLeuLysAspSerGluLysLysArgTyrLeuGlyLeu 308
Dy AGCTACGGGTACCTGGCGGATGTGGCAGCCAGAGCGAGAACTACCGCTGATGAGCGCGAA 508
Qy 309 AlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThr 328
Dy CGCCGGTACGAGTACAGTGGCGTCAAGGCCCTTCTGTAATAATCGCGCTATGAGCGCGAA 448
Qy 329 ValSerPheLeu-----ProAlaGlnHisThrValGly 339
Dy 447 CTGAGATGTTAGAGAGCCCGATCTTCTACTGACCAGCGCGCTGGAGGACATTCGCGAG 388
Qy 340 SerProArgAsp-----ArgLysProCysArgAlaGlyCys 351
Dy AGTCGGATAGTGTGTGCTGCTGGGAGAGTCAGTGCCATCCGTCTGTCTATGCCAATTGC 328
Qy 352 PheValCys-----ArgGlnSerLysGlnGlnLeuGlu 363
Dy 327 CAGCGCTCAGCTTCGCCAGCAGCATACAGGAACAGCGATCTCATTTGTTTCATCCAGAG 268
Qy 364 GluGlnLysLysAlaLeuTyrGlyLeuGluAlaGluAspValGlu----- 379
Dy 267 GAATCTAAAGGCCA-----GAGCGCATACAGGAGGTAGAAACAGAGGACTCT 220
Qy 379 ----- 379
Dy 219 CATCTAGCGCGCAGTGAAGCAGCACTCTTAGCGCTCTGCGCGTCCAGGCAATCTTCGA 160
Qy 380 -----GluTyrGlnVal 384
Dy 159 TTGCCCACTGGCTCCATTTTCATCAATAGGAACCTCGGCAACGATCAGTGGAGGTTGTG 100

Alignment Scores:
Pred. No.: 2,34e-24 Length: 901
Score: 302.50 Matches: 85
Percent Similarity: 41.24% Conservative: 35
Best Local Similarity: 29.21% Mismatches: 82
Query Match: 10.47% Indels: 89
DB: 4 Gaps: 9

US-10-631-958-10 (1-537) x US-09-270-767-30448 (1-901)

Qy 317 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu----- 332
Dy 2 AAGSCCTTCTGAATAATCGCGCTATGACCGCGAATGTAGAGAGCCCGAT 61
Qy 333 -----ProAlaGlnHisThrValGlySerProArgAsp----- 343
Dy 62 CTTCTACTGACACCGCGCTGAGGACATTCGCGAGATCCGAGATCGGATAGTGTGTCGCTG 121
Qy 344 -----ArgLysProCysArgAlaGlyCysPheValCys----- 354
Dy 122 GGAGAGTCAGTGCATCCGTCTGCTATGCCAATGCGCGCTGCGCTCGCAGCAGC 181
Qy 355 -----ArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrGly 371
Dy 182 ATACAGGACAGCGATCTCTATTGTTTCATCCAGAGGAATCTAAAGAGCA----- 232
Qy 372 LeuGluAlaAlaGluAspValGlu----- 379
Dy 233 ---GAGCGCAATCAGAGGTAGAAACAGAGACTCTCATCTAGCGCGAGTGAAGCAGCA 289
Qy 379 ----- 379
Dy 290 CTCCTGAGGCTGCTCGCGTCCAGGCAATCTTCGATTGCCCTCCATTCATCA 349
Qy 380 -----GluTyrGlnValValCysGlyLysPheLeuAlaIleAsn 392
Dy 350 ATGAGGAACCTCGGCAACGATCAGTGAAGGTTGTGGGGGCAATTTCTTTATGATCTGC 409
Qy 393 AlaThrAsnMetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHis 412
Dy 410 GGCCGCAACATAACCTCGCTGCGCAGGAGTCCCAATGGCATCTCCGTTACAGTCAT 469
Qy 413 LeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeu 432
Dy 470 CTGGGTATGGTTCCTGGACCTGATTTCTGTGTAAGAGACCTCAGCTGCTCAACACGTG 529
Qy 433 ArgPheLeuIleArgHisThrAsnGlnGln---AspGlnPheAspPheThrPheValGlu 451
```

Db 530 CGTTTCCTCAACACGGCGGCGGAGTGGTATTCGCAATTTGCTTTTGAGAG 589
QY 452 ValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluSerAsp 471
Db 590 GTATATCGCACAGGAAATTCGGTTTCAGAACTTTTTCGCCAGCGAGGAGCTACAGC 649
QY 472 LeuLysGluGlyGlyLysLysArgPheGlyHisIleCys----- 484
Db 650 TTGGCAGGGTCT-----TGTCAGCCGATTAACACGGCTGAA 685
QY 485 -----SerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCysAsp 502
Db 686 GAAATGACCGCCCACTCTCC-----TCGACGGAGTTTTCAGCTGGAACTGTGAT 736
QY 503 GlyGluValLeuHisSerProLalleGluValArgValHisCysGlnLeuValArgLeu 522
Db 737 GGAGAGGTGCTGACCGACCTGGACATAACCATGCGATGCAATTGTGAGCTCATCGAGGTC 796
QY 523 PheAlaArgGlyIleGluGluAsnProLysPro 533
Db 797 TTTCATGCGGGGTCCCAATTCCTACAGCAAGCCC 829

RESULT 5

US-09-970-516-3
; Sequence 3, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1857)
; OTHER INFORMATION:
US-09-970-516-3

Alignment Scores:

Pred. No.: 5,08e-21 Length: 1857
Score: 278.00 Matches: 91
Percent Similarity: 44.24% Conservative: 55
Best Local Similarity: 27.58% Mismatches: 155
Query Match: 9.63% Indels: 30
DB: 4 Gaps: 6

US-10-631-958-10 (1-537) x US-09-970-516-3 (1-1857)

QY 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTTrpTrpArgSerProGlyPro 40
Db 128 TCGGCCCAACCTGAAGCCAGGCCCGCGG---GTGGCTGTGCTCCGTTGCCGAGGTCT 184
QY 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 60
Db 185 CAGGCTGTGTGACCC---TGCGAAGCCGCGAGCCCTCAGACTCAGCGGCTACTTCTGCA 241
QY 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 242 TCTAC-ACCTACCTCGGGCGCGCGGGCGCGGCGGCGAGCCACTCGCACCTTCGG 300
QY 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100
Db 301 GCAGATGGGGCGCCCACTACGAGAGAACCGTGGCGAGGCCCGCCAGCGCTGGGCCACTGCC 360
QY 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120

Db 361 CTCACCTGTCTG-----CTCCGAGACTGCCACTGCCCGGGGATGGG 402
QY 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 403 GAGATCACCCCTGACCTGTACTACCTCGGCGCGCCCGTTGCTTCTATTGGTCAATCCCTTT 462
QY 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 463 GGGGCTCGGGGCTGGCCCTGGCAGTGGTGTAAAGAACCAACCATGCTTCCCATGATCTCTGAA 522
QY 161 AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 523 GCTGGCTGTCTTCACTCATCCAGACAGACAGAACCAACCGCGGAGGTGGTTC 582
QY 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 583 CAGGGCTGAGCCTGAGTGAGTGGATGCGATCGTCACGGTCTCGGAGAGCGGCTGCTC 642
QY 201 SerGluValLeuHisGlyLeuIleGlyThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 643 CATGAGTGTCTGAACGGGCTCTAGATGCG-----GTGGGATCTCTCCCTGCGGCTG 678
QY 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 679 TGGGAGGAAGCTGTGAAGATGCT-----GTGGGATCTCTCCCTGCGGCTG 726
QY 241 ThrAspCysValCysTyrSer-----ThrValGlyThr 251
Db 727 GGCAACGGCTGGCGGAGCAGTGAACAGCAGCGGGGATTTGAGCCAGCCCTGGGCTC 786
QY 252 SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal 271
Db 787 GACCTGTGTCTCAACTGCTACTGTGCTGTCGGGGTGGTGGCCACCCACCTGACCTG 846
QY 272 SerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGly 291
Db 847 CTCTCCGTGACGCTGGGCTCGGCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 906
QY 292 PheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyr 311
Db 907 TTCGTGTGATGTGGATATCCAGAGCGAGCGCTTCAGGGCTTGGGCGAGTGGCCGCTC 966
QY 312 AspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPhe 331
Db 967 ACACCTGGGCACGGTGTGGGCTCGCCACACTGACACACTACCGCGGAGCGCTCTCTCTAC 1026
QY 332 LeuProAlaGlnHisThrValGlySerPro 341
Db 1027 CTCCCGCCCACTGTGGAACCTGCTCGGCC 1056

RESULT 6

US-09-817-676A-13
; Sequence 13, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 13
; LENGTH: 2380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

```
; LOCATION: (7)..(1960)
; PUBLIC INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245447
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-13

Alignment Scores:
Pred. No.: 7,78e-21 Length: 2380
Score: 278.00 Matches: 91
Percent Similarity: 44.24% Conservativeness: 55
Best Local Similarity: 27.58% Mismatches: 155
Query Match: 9.63% Indels: 30
DB: 4 Gaps: 6

US-10-631-958-10 (1-537) x US-09-817-676A-13 (1-2380)

Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTrpArgSerProGlyPro 40
Db 134 TCGCGCCCAACCTGAAGCCAGCCGCCCGG---GTGGCTCGTCCGCTGGCGGAGGTCT 190
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db 191 CAGGCTGTGCACCC---TGGGAAGCCGACCCCTCAGATCAGCGGCCCTACTTCTGCA 247
Qy 61 GluGluThrAspValHisGlyHisGlnGlySerGlyLysTyrGlnLysMetGluLys 80
Db 248 TCTAC-ACCTACCTCGGGCGGGCGGGCGGGCCGAGCCGACCTCGACCTTCGG 306
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTyrLysTrpAla 100
Db 307 GCAGATGGGGCGCCACCTACGAGAACCGTGCAGAGCCGAGCCGAGCGCTGGGCCACTGCC 366
Qy 101 GlnValThrPheTyrProGluGlnGlnCysHisLeuTyrLeuGlnThrLeuArg 120
Db 367 CTCACCTGCTG-----CTCGAGGAGCTGCCACTCCCGGGGATGG 408
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 409 GAGATCACCCCTGACCTGCTACCTCGGCGCGCCGTTGCTTCTATTGGTCAATCCCTTT 468
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 469 GGGGGTCGGGCGCTGGCTGGCAGTGGTGAAGAACCAACCGTCTCCCATGATCTCTGAA 528
Qy 161 AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 529 GCTGGGCTGCTCTCAACCTCATCAGACAGAACAGACCCAGCCCGCGGAGCTGGTC 588
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 589 CAGGGCTGAGCCTGAGTGAGTGGATGCGTCACGGTCTCGGAGAGCGGGCTGCTC 648
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 649 CATGAGTGTGAACGGGCTCTAGATGCC-----GTGGGATCTCTCCCTGGCGGCTCG 684
Qy 221 HisProArgAlaValLeuValProSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 685 TGGGAGGAAGCTGTGAAGATGCT-----GTGGGATCTCTCCCTGGCGGCTCG 732
Qy 241 ThrAspCysValCysTyrSer-----ThrValGlyThr 251
Db 733 GCACAGCGCTGGCGGAGCAGTGAACAGCAGCGGGGATTTGACCCAGCCCTGGGCTC 792
Qy 252 SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal 271
Db 913 GACCTGTTGTCAACTGCTCACTGTTGCTGTGCGGGGTGGTGCCACCCACTGGACCTG 852
Qy 272 SerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGly 291
Db 853 CTCTCCGTGAGCTGGGCTCGGGCTCCGCTGTTCTCTCTTCTGCTGTGTGGCTGGGC 912
Qy 292 PheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyr 311
Db 913 TTCGTGTGAGATGGATATCCAGAGCGGCTTCAGGGCCTTGGGCGAGTGGCCGCTTC 972
Qy 312 AspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPhe 331
Db 973 ACCTGGGCGAGCTGGGCGCTCGCCACACTACCGCGGAGCGCTCTCTCTAC 1032
Qy 332 LeuProAlaGlnHisThrValGlySerPro 341
Db 1033 CTCCCCGCCACTGTGGAACCTGCTGCC 1062

RESULT 7
US-09-817-676A-11
; Sequence 11, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; TITLE OF INVENTION: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; PRIOR FILING DATE: 2001-03-26
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2698
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (387)..(2237)
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245448
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-11

Alignment Scores:
Pred. No.: 1.03e-19 Length: 2698
Score: 269.00 Matches: 98
Percent Similarity: 41.64% Conservativeness: 59
Best Local Similarity: 25.99% Mismatches: 159
Query Match: 9.31% Indels: 62
DB: 4 Gaps: 8

US-10-631-958-10 (1-537) x US-09-817-676A-11 (1-2698)

Qy 19 GlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTrpArgSerPro 38
Db 427 CAATCTGTACCGCGAGTTTGGTTCTTCTACCGGCCAACGGCCCGCTTCCCTCACCC 486
Qy 39 Gly-----ProGlyAlaGlyAlaProGlyAlaAsp 48
Db 487 TCACAACACAGCCCTACACATACAGCGACTACGCCCAAGCCAGAGAGCCCGGCCGAG 546
Qy 49 AlaCysSerValProValSerGluIleIle-----AlaVal 60
Db 49 AlaCysSerValProValSerGluIleIle-----AlaVal 60
```

| | | | | | | | | | | | | | | | | | | | | | | |
|------|-------|---------|--------------------|-----------------------------------------------|-----------------------------------------------------|------------------------------|--------------------------------------------|-----|-----|-----|-----|-----|-----|-------|-------|-------|-------|-------|-------|-------|-----|-----|
| 547 | ATG | TCTA | -GTC | TCTCTCTGGATGAGGTC | TCTGGGCGTGTGGCACCCTGTGCAGAGCGCGTAGCCCC | 605 | | | | | | | | | | | | | | | | |
| 61 | Gluc | Thr | Asp | Val | His | 72 | | | | | | | | | | | | | | | | |
| 606 | GAG | GAC | ACT | TGC | GCGCTACTTCTTCGTACTTACACCTACCCACCGTGGCCGTCGAGGGGGCGG | 665 | | | | | | | | | | | | | | | | |
| 73 | Gly | Leu | Trp | Gln | Leu | Met | Glu | Leu | Pro | Tyr | Ala | Phe | Thr | Val | His | Cys | Val | Leu | Arg | Ala | 92 | |
| 666 | CGC | AG | ACT | ACG | CGG | AGCTTC | TCCGGCGCGATGGGGCCACCACTTATGAGGAGAAATCGTGCA | 725 | | | | | | | | | | | | | | |
| 93 | Arg | Arg | His | Arg | Tyr | Leu | Val | Thr | Ala | Gln | Val | Thr | Phe | Trp | Cys | Pro | Glu | Glu | Gln | Leu | Cys | 112 |
| 726 | GAG | CC | CAG | CGCTGGG | CCACTGCCCTCACGTGCTC | 767 | | | | | | | | | | | | | | | | |
| 113 | His | Leu | Trp | Leu | Gln | Thr | Leu | Arg | Glu | Met | Leu | Glu | Leu | Leu | Leu | Ser | Arg | Pro | Leu | His | 132 | |
| 768 | GG | AGT | GCTCTGT | CAGG | GGACAGGAAATCA | CCCTCGAATTTGCTGCCCGGAAGCCAGG | 822 | | | | | | | | | | | | | | | |
| 133 | Leu | Leu | Val | Phe | Leu | Asn | Pro | Phe | Gly | Gly | Gln | Gly | Leu | Leu | Leu | Arg | Gly | Thr | Gln | Arg | 152 | |
| 828 | CTG | CTCATAT | TGGTCAAT | TCCCTTTGGGGGCGGGGCTGGCGCTCGCAGCGCTGTATGGAC | 887 | | | | | | | | | | | | | | | | | |
| 153 | Lys | Val | Ala | Pro | Leu | Phe | Thr | Leu | Ala | Ser | Leu | Thr | Thr | Asp | Leu | Val | Thr | Glu | His | 172 | | |
| 888 | CAC | GTGGTGC | CAATGATCTCTG | AGCTGGGCTGTCTTCAACCTCATACAGACAGACGA | 947 | | | | | | | | | | | | | | | | | |
| 173 | Ala | Asn | Gln | Ala | Lys | Glu | Thr | Leu | Tyr | Glu | Leu | Leu | Asn | Leu | Asp | Lys | Tyr | Asp | Gly | Val | 192 | |
| 948 | CAG | AAC | CAATGCC | GTGAGCTGGT | CGAGGGTTAAGCCCTGAGTGGTGGGAAGGCATTGTC | 1007 | | | | | | | | | | | | | | | | |
| 193 | Cys | Val | Gly | Gly | Asp | Gly | Met | Phe | Ser | Glu | Val | Leu | His | Gly | Leu | Leu | Gly | Arg | Thr | Gln | 212 | |
| 1008 | ACT | GTGTCT | CGAGACGGGCTGCTTTAC | GAGTGTGTAATGGGCTCTTGATCGG | 1061 | | | | | | | | | | | | | | | | | |
| 213 | Arg | Ser | Ala | Gly | Val | Asp | Gln | Asn | His | Pro | Arg | Ala | Val | Leu | Val | Pro | Ser | Ser | Leu | Arg | 232 | |
| 1062 | ----- | ----- | ----- | CCAG | ACTGGGAGGATCCGCTGGGATGCC | 1091 | | | | | | | | | | | | | | | | |
| 233 | Ile | Gly | Ile | Leu | Pro | Ala | Gly | Ser | Thr | Asp | Cys | Val | Cys | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 245 | |
| 1092 | ATT | GGTGTCT | CCCTCGGATCGG | CAATGCGCTAGCTGGGCGGTGAGCCATCATGGC | 1151 | | | | | | | | | | | | | | | | | |
| 246 | ----- | ----- | ----- | Tyr | Ser | Thr | Val | Gly | Thr | Ser | Asp | Ala | Glu | Thr | Ser | Ala | Leu | His | Ile | Val | 263 | |
| 1152 | GGG | TTTGAG | CAGGTTGTC | GGTGTGACCTGTGCTCAACTGCTCGCTTCTTCTCTCGCGT | 1211 | | | | | | | | | | | | | | | | | |
| 264 | Gly | Asp | Ser | Leu | Ala | Met | Asp | Val | Ser | Ser | Val | His | His | Asn | Ser | Thr | Leu | Leu | Arg | Tyr | 283 | |
| 1212 | GGT | GGCAG | CCATCTCT | CTGGACTTGCTCTCTGTGACGCTAGCCTCGGATCCCGCTGTTT | 1271 | | | | | | | | | | | | | | | | | |
| 284 | Ser | Val | Ser | Leu | Leu | Gly | Tyr | Gly | Phe | Tyr | Gly | Asp | Ile | Leu | Ser | Asp | Ser | Glu | Leu | Lys | 303 | |
| 1272 | TC | TTCTCT | GTGTCAGTGGGCTCGG | GATCTTGTGCAGATGGGACATTCACAGTGGCGCTC | 1331 | | | | | | | | | | | | | | | | | |
| 304 | Arg | Trp | Leu | Gly | Leu | Ala | Arg | Tyr | Asp | Phe | Ser | Gly | Leu | Lys | Thr | Phe | Leu | Ser | His | His | 323 | |
| 1332 | AGG | CCCTGG | CGACGCTCG | ATTCCACTACCTGGGTGCAGTGTGCTAGCGCTGGCGCTCGTTCAT | 1391 | | | | | | | | | | | | | | | | | |
| 324 | Cys | Tyr | Glu | Gly | Thr | Val | Ser | Phe | Leu | Pro | Ala | Gln | His | Thr | Val | Gly | Ser | Pro | Arg | As | 343 | |
| 1392 | AC | TAC | CGTGG | CGCCTCTCT | CACTCCCGCT | 1430 | | | | | | | | | | | | | | | | |
| 343 | par | Gly | Pro | Cys | Arg | Ala | Gly | Cys | Phe | Val | Cys | Arg | Gln | Ser | Lys | Gln | 359 | | | | | |
| 1431 | GA | ACC | AGCTT | GGCCATCC | CCAGGCCACAGTCTGCTCGAGCAAGTCAG | 1479 | | | | | | | | | | | | | | | | |

```

? APPLICANT: Pu, Xia
? APPLICANT: Richard, D'ANDREA J
? APPLICANT: Jennifer, BAMBLE R
? APPLICANT: Mathew, VADAS A
? TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
? FILE REFERENCE: PITSON=1
? CURRENT APPLICATION NUMBER: US/09/959,897
? CURRENT FILING DATE: 2001-11-13
? PRIOR APPLICATION NUMBER: PCT/AU00/00457
? PRIOR FILING DATE: 2000-05-12
? PRIOR APPLICATION NUMBER: AU PQ 0339
? PRIOR FILING DATE: 1999-05-13
? PRIOR APPLICATION NUMBER: AU PQ 1504
? PRIOR FILING DATE: 1999-07-08
? NUMBER OF SEQ ID NOS: 56
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 1
? LENGTH: 1205
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (33)..(1184)
? OTHER INFORMATION:
? US-09-959-897-1

```

```
Db 573 AGTGAGAGTATCGCGTCTGGGCGAGATGCGCTTCACTCTGGGCACTTTCTCGCTG 632
Qy 320 LeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGly 339
Db 633 GCACCTTGGCGACTTACCGCGCGAGTGGCTTACCTCCCTGTAGGAAGA---GTGGT 689
Qy 340 SerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGln 359
Db 690 TCCAAGACACCTGCTCCCCC-----GTTGGTGGTCCAGAGGCGCCGGTA 734
Qy 360 GlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal--- 378
Db 735 GATGCACAC-----CTTGTGCCACTGGAG-----GAGCCAGTGC 770
Qy 379 GluGluTrpGlnValValCysGlyLys-----PheLeuAlaAlaAsnAlaThr 394
Db 771 TCTCAGTGCAGAGTGGTCCCGCGAGGACTTTGTGTAGTCTCTGGGCACTGCTGCATCG 830
Qy 395 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 414
Db 831 CACTGGGCACTGAGATGTTGCTGCACCCATGCGCGCTGTGCAGCT----- 878
Qy 415 AspGlySerSerAspLeuLeuLeuLeuArgLys---CysSerArgPheAsnPheLeuArg 433
Db 879 ---GGCGTCATGCTCTGCTACGTGCGGCGGAGTGTCTGCTGCATGCTGCTGCGC 935
Qy 434 ---PheLeuLeuLeuArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluVal 452
Db 936 CTCCTTCCTGGCCATGGAGAGGCGCATATGAGTATGAATCCCTACTTCTGGTATAT 995
Qy 453 TyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu 472
Db 996 GTGCCGTGTCGCTTCCG-----TTGGAGCCCAAGAT----- 1031
Qy 473 LysGluGlyGlyLysLysArgPheGlyHisLeuCysSerSerHisProSerCysCys 492
Db 1032 -----GGGAAAGGTATGTTGCA----- 1049
Qy 493 ThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaLeuGlu 512
Db 1050 -----GTGGATGGGAAATTGATGGTTAGCGAGCGCGTCAG 1085
Qy 513 ValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnPro--- 531
Db 1086 GGCAGGTGCACCCAACTACTTCTGGATGTCAGCGGTTCCGTGGAGCCCGCCCGCAGC 1145
Qy 532 ---LysPro 533
Db 1146 TGAAGCCC 1154

RESULT 9
US-09-970-516-1
; Sequence 1, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1155)
; OTHER INFORMATION:
US-09-970-516-1
```

```
Alignment Scores:
Pred. No.: 7,83e-20 Length: 1155
Score: 264.50 Matches: 107
Percent Similarity: 41.37% Conservatives: 68
Best Local Similarity: 25.30% Mismatches: 169
Query Match: 9.16% Indels: 79
DB: 4 Gaps: 16

US-10-631-958-10 (1-537) x US-09-970-516-1 (1-1155)
Qy 129 ArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlnGlyLysArg 148
Db 37 CGGCCCTCGCGCTGTGTGTGTGTGAACCGCGCGCGCAAGGCGCTTCGAG 96
Qy 149 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIle 168
Db 97 CTCTTCGGAGTACGCTGCAGCCCCCTTTGGCTGAGGCTGAATCTCTTCAGCTCATG 156
Qy 169 ValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyr 188
Db 157 CTCACTGAGCGCGGAACACACGCGCGGAGCTGGTGGCTCGGAGGAGCTGGGCGCTGG 216
Qy 189 AspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeu 208
Db 217 GACGCTCTGTGTGTGTGTCTGTGAGACGCGCTGATGCAGAGTGTGTAACGCGCTCATG 276
Qy 209 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 228
Db 277 GAGCGGCGCTGACTGGGAGACCGCCATCCAGAAG----- 309
Qy 229 SerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCys----- 245
Db 310 -----CCCTGTGTAGCTTCCAGCAGGCTCTGGCAACGCGCTGGCAGCTTCCTTG 360
Qy 246 -----TyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 361 AACCATTTATGCTGTGTATGAGAGGTCAACATGAAGACCTCTCTGACCACTGCACGCTA 420
Qy 261 IleValValGlyAspSerLeuAla---MetAspValSerSerValHisHisAsnSerThr 279
Db 421 TTGCTGTGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Qy 280 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuLysAsp 299
Db 481 CTGGCGCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Qy 300 SerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 319
Db 541 AGTGAGAAATATCGCGCTGTGGGAGATGCGCTTCACTCTGGGCACTTCTCTCGCTCTG 600
Qy 320 LeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGly 339
Db 601 GCAGCCTGCGCACCTACCGCGCGCGAGTGGCTTACCTCTCTGTAGGAAGA---GTGGGT 657
Qy 340 SerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGln 359
Db 658 TCCAAGACACCTGCTCCCCC-----GTTGTGGTCCAGAGGCGCCCGGTA 702
Qy 360 GlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal--- 378
Db 703 GATGCACAC-----CTTGTGCCACTGGAG-----GAGCCAGTGC 738
Qy 379 GluGluTrpGlnValValCysGlyLys-----PheLeuAlaAlaAsnAlaThr 394
Db 739 TCTCAGTGCAGAGTGGTGGCGCGAGGAGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 798
Qy 395 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 414
Db 799 CACTGGGCACTGAGATGTTGTGTGACCCCATGGCGCGCTGTGCGAGCT----- 846
Qy 415 AspGlySerSerAspLeuLeuLeuLeuArgLys---CysSerArgPheAsnPheLeuArg 433
```

```
Db      847 ---GGCGTCATCATCTGTTCTACGTGCGGGCGGAGTGTCTCGTGCCTATGCTGCTGCGC 903
Qy      434 ---PheLeuIleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluVal 452
Db      904 CTCCTCTCGCCATGCAGAGGCGCAGGATATGAGTATGAATGCCCTACTTGGTATAT 963
Qy      453 TyrArgValLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu 472
Db      964 GTGCCCGTGTGCTCTCCGC-----TTGGAGCCCAAGAT----- 999
Qy      473 LysGluGlyGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCys 492
Db      1000 -----GGAAAGGTGTGTTGCA----- 1017
Qy      493 ThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGlu 512
Db      1018 -----GTGGATGGGAATTGATGTTAGCGAGCGCGTGCAG 1053
Qy      513 ValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnPro--- 531
Db      1054 GGCCAGGTGCACCCAACTACTTCTGGATGTGTACGGTTCGTGGAGCCGCCCGCCAGC 1113
Qy      532 ---LysPro 533
Db      1114 TGAAGCCC 1122

RESULT 10
US-09-949-016-1155
; Sequence 1155, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1155

Alignment Scores:
Pred. No.:      2,15e-19      Length:      1783
Score:          263.50      Matches:     107
Percent Similarity: 41.22%      Conservative: 69
Best Local Similarity: 25.06%      Mismatches:  164
Query Match:      9.12%      Indels:      87
DB:              4          Gaps:         17

US-10-631-958-10 (1-537) x US-09-949-016-1155 (1-1783)

Qy      129 ArgProLysHisLeuLeuValPheIleAsnPropheGlyGlyLysGlyGlnGlyLysArg 148
Db      373 CGGCCCTGCGCGTGTGTGTCTGAACCGCGCGCGCAAGGCAAGCCCTTGCAG 432
Qy      149 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIle 168
Db      433 CTCCTCCGAGACGACGTGCGAGCCCTTTTGGCTGAGGCTGAAATCTCCTTCACGCTGAT 492
Qy      169 ValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyr 188
Db      493 CTCACGTAGCGCGGAACACACGCGGGAGCTGGTGTGGAGGAGCTGGCGCGCTGG 552
```

```
Qy      189 AspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIle 208
Db      553 GACCTCTGGTGGTCATCTCTGGAGACGGGCTGATGCACGAGGTGGTGAACGGGCTCATG 612
Qy      209 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 228
Db      613 GAGCGGCTGACTGGGAGACCGCATCCAGAG----- 645
Qy      229 SerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCys----- 245
Db      646 -----CCCTGCTGTAGCTCCAGCAGGCTCTGGCAACGGCTGGCAGCTTCCTTG 696
Qy      246 -----TyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db      697 AACCATATGCTGCTATGACAGGTCAACATGAAGACCTCTCGACCACTGCACGCTA 756
Qy      261 IleValValGlyAspSerLeuAla---MetAspValSerSerValHisHisAsnSerThr 279
Db      757 TTGCTGTGCCCGCGCTGCTGCACCCATGACCTGTCTCTGCACACGGCTTCGGGG 816
Qy      280 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAsp 299
Db      817 CTGCGCTCTTCTCTGCTCAGCTGCGCTGGCGCTTCATTGCTGATGTGCACCTAGAG 876
Qy      300 SerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 319
Db      877 AGTGAGAAGTATCGCGCTCTGGGGGAGATGCGCTTCACTCTGGGCACCTTCCTCGCTG 936
Qy      320 LeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGly 339
Db      937 GCAGCCCTGGCACCTACCGCGCGGAGTGGCTTACCTCCTCTGTAGGAGAG---GTGGGT 993
Qy      340 SerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGln 359
Db      994 TCCAAGACACCTGCTCCCCC-----GTTGTGGTCCAGCAGGCGCCGCTA 1038
Qy      360 GlnLeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal--- 378
Db      1039 GATGCACAC-----CTTGTGCCACTGGAG-----GAGCCAGTGCCTC 1074
Qy      379 GluGluTrpGlnValValCysGlyLys-----PheLeuAlaIleAsnAlaThr 394
Db      1075 TCTCAGCTGGACATGGTGGCGGACGAGACTTTGTGTATCTCTGCTACTGCTGCACCTG 1134
Qy      395 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 414
Db      1135 CACCTGGCGAGTGAGATGTTTGTGTCACCCATGGCGCGCTGTGCAGCT----- 1182
Qy      415 AspGlySerSerAspLeuIleLeuLysLysLysArgLys---CysSerArgPheAsnPheLeuArg 433
Db      1183 ---GGCGTCATGCATCTGTTCTACGTGCGGCGGAGTGTCTCGTGCCATCTGCTGCGC 1239
Qy      434 PheLeuIle-----ArgHisThrAsnGlnAspGlnPheAspPheThr 448
Db      1240 TTCTTCTGCGCATGGAGAGGCGGAGCATATG-----GAGTATGAATGCCCC 1287
Qy      449 PheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGlu 468
Db      1288 TACTTGGTATATGTCGCGTGGTGCCTTCGC-----TTGGAGCCCAAG 1332
Qy      469 AspSerAspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHisPro 488
Db      1333 GAT-----GGGAAAGGTGTGTGCA----- 1353
Qy      489 SerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSer 508
Db      1354 -----GTGGATGGGAATTGATGTTAGC 1377
Qy      509 ProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgIleGlu 528
Db      1378 GAGCGCTGCGAGGCCAGGTGCACCCAACTACTTCTGATGGTGCAGTGGTGGCTGGAG 1437
Qy      529 GluAsnPro-----LysPro 533
```


QY 337 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 356
 Db ||||| : : : : :
 643 ---GTAGAACTGTGGCCCTCTAAGAGACCCGCC----- 672
 QY 357 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaGlu 376
 Db ||||| : : : : :
 673 ---TCTACACTGGTGCAGAACAGGGCCCGCTGCACACACACTTGTCTCTCGAGGAG 726
 QY 377 AspVal---GluGluTyrGlnValValCysGlyLys-----PheLeuAlaile 391
 Db ||||| : : : : :
 727 CCAGTGCCTCTCATTTGACCTGTGTACCAAGACAGGACTTTGTCTGTGCTGTGCTG 786
 QY 392 AsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAla 411
 Db ||||| : : : : :
 787 CTACACACCCACCTGAGCTCCGAGCTGTTTCAGCACCCATGGCGCGCTGTGAGGCT--- 843
 QY 412 HisLeuGlyAspGlySerSerAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 430
 Db ||||| : : : : :
 844 -----GGTGTATTGTCATCTGTTTACGTACGTGCGGGGGTGTCAAGGGCTGCG 891
 QY 431 PheLeuArg---PheLeuLeuArgHisThrAsnGlnGlnAspGlnPheAspPheThrPhe 449
 Db ||||| : : : : :
 892 CTGCTGCCCTCTCTCTGCCCATGCAGAAAGGCGACATATGGAACCTTGATCTGTCCTAC 951
 QY 450 ValGluValTyrArgValLysPheGlnPheThrSerLysHisMetGluAspGluAsp 469
 Db ||||| : : : : :
 952 CTGGTTTCATGCGCGTGTGCTTTCGCGCTGGAGCCAGGAGCGGCGTGTGTT 1011
 QY 470 SerAspLeuLysGluGlyLysLysArgPheGlyHisLysCysSerSerHisProSer 489
 Db ||||| : : : : :
 1012 TCT-----GTGGATGAGGGCTGATGG--TATGTGAAGCTGTGACGGGCCAAG----- 1057
 QY 490 CysCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerPro 509
 Db ||||| : : : : :
 1058 -----TGCACCCAACTACCTTTGGATGGTGTGTGGCAGCAGAGATGCCCAT---CCG 1108
 QY 510 Ala 510
 Db ||||| : : : : :
 1109 GCC 1111

RESULT 13

US-09-248-796A-1756
 ; Sequence 1756, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 1756
 ; LENGTH: 1611
 ; TYPE: DNA
 ; ORGANISM: Candida albicans
 US-09-248-796A-1756

Alignment Scores:
 Pred. No.: 2,04e-15 Length: 1611
 Score: 228.00 Matches: 81
 Percent Similarity: 42.86% Conservative: 57
 Best Local Similarity: 25.16% Mismatches: 143
 Query Match: 7.89% Indels: 42
 DB: 4 Gaps: 9

US-10-631-958-10 (1-537) x US-09-248-796A-1756 (1-1611)

QY 119 LeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsn 138

Db ||||| : : : : :
 454 TTAACAGAACTCATATGGGAAAAGTATAATTCGGCCCTCT---ATTTTGGTATTGATAAAC 510
 QY 139 ProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPhe 158
 Db ||||| : : : : :
 511 CCGCATGGCGCCCAAGACACGCCAAACTATTTTACAAAATAAAATCTTTACCAATAATTA 570
 QY 159 ThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGlu 178
 Db ||||| : : : : :
 571 CAGCGCGCTCGTCTAATGTTACGTATTTTGAACCTAAATATCATGACACGCCACTGAG 630
 QY 179 ThrLeuTyrGluLeuAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 198
 Db ||||| : : : : :
 631 ATTGCGCGCTGAGCTAGATGCAATGATTATGATAATAATGTTGTTGTTCTTGGCGATGG 690
 QY 199 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 218
 Db ||||| : : : : :
 691 ATACCTCATGAAGTATCAATGGCTTCTATCTTCCTCCAGAT----- 732
 QY 219 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 238
 Db ||||| : : : : :
 733 -----AAAGGTTTACTGGCATTCAACAAAATTCGAGTTACTCAATTACCTTGT 780
 QY 239 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 258
 Db ||||| : : : : :
 781 GGGTCAGGGAACCGGTTGAGTTGAGTACACATGTTAGTAAATAATGCTTCAGTTGCAACT 840
 QY 259 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHis----- 276
 Db ||||| : : : : :
 841 CTTTATATGTTGAAGCTCATAGACAAAATTTGATGCTTATACCCAAAGGTACA 900
 QY 277 ---AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp 295
 Db ||||| : : : : :
 901 GGAAGTGAAGAATAACGAAATTTGCTATTTTAAAGTCAGTCTACGGTATTATTTGCTGAT 960
 QY 296 IleIleLysAspSerGluLysLysArgTrpLeuAlaArgTrpLeuAlaArgTyrAsp----- 312
 Db ||||| : : : : :
 961 TCTGATATTGGAACAGAACATTTACGTTGGTTCCTATACGGTTTGAACCTGGAGTG 1020
 QY 313 -----PheSerGlyLeuLys-----ThrPheLeuSerHisHisCys 324
 Db ||||| : : : : :
 1021 ATACAAAAGTGTCTCGGGGCAAAATATCTTTGTGACCTATTTGTGAAATACAAG--- 1077
 QY 325 TyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArg 344
 Db ||||| : : : : :
 1078 TAGATAATAATTGAGATTTTGAATCATGTAATGATTACTTGTGATAATATGAT--- 1134
 QY 345 LysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlu 364
 Db ||||| : : : : :
 1135 -----ACTGAAAACGAGTTACCCATTGTCACCTGAA 1164
 QY 365 GlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal---GluGluTyrGln-Va 383
 Db ||||| : : : : :
 1165 GAGAACTTGCAAAATAACGACCCCTGATTTAGATCAACCTGTCCTCCCAATGATGGAGACAT 1224
 QY 383 lValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgSe 403
 Db ||||| : : : : :
 1225 ATTCTCAAGAAATTTCTATAATTTGAACATCTTGTATGTTGGCAAAATGCCATTGCTG 1284
 QY 403 rProArg-----GlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIle 421
 Db ||||| : : : : :
 1285 TCTGCAGATACTCAGTGTTCCTCTACCGAAGCATGGTTCAATGGACATGAT 1343
 QY 421 eLeu 422
 Db ||||| : : : : :
 1344 TGTC 1347
 RESULT 14
 US-09-614-221A-399
 ; Sequence 399, Application US/09614221A
 ; Patent No. 6723837
 ; GENERAL INFORMATION:
 ; APPLICANT: Karunanandaa, Balasulojini

Qy 352 PheValCysArgGlnSerLysGlnGlnLeuGlu 362
 Db 1258 AAATATGCTGCCAAATCTAAAGGAATTAAAA 1290

RESULT 15
 US-09-949-016-3940
 ; Sequence 3940, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0
 ; SEQ ID NO 3940
 ; LENGTH: 1764
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-3940

Alignment Scores:
 Pred. No.: 1,5e-12 Length: 1764
 Score: 203.50 Matches: 101
 Percent Similarity: 40.48% Conservative: 67
 Best Local Similarity: 24.34% Mismatches: 167
 Query Match: 7.05% Indels: 81
 DB: 4 Gaps: 17

US-10-631-958-10 (1-537) x US-09-949-016-3940 (1-1764)

Qy 137 IleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaPro 156
 Db 397 CTGAACCCG--CGAAGCGAAGGCAAGGCTT-CAGCTCTCCGGAGTACGTCGACGCC 452

Qy 157 LeuPheThrLeuAlaSerIleThrAspIleIleValThrGluHisAlaAsnGlnAla 176
 Db 453 CTTTGTGCTGAGGCTGAATCTCTTTCACGCTGATGCTCACTGAGCGCGCGAACCACGCG 512

Qy 177 LysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGly 196
 Db 513 CGAGACTGATGCATCGGAGAGCTGGCGCGCTAGGACGCTCTGGTGGTCATGTCGTGA 572

Qy 197 AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly 216
 Db 573 GACGGGCTGATCAGCAGGTTGGTGAACGGGCTCATGGAGCGGCTGACTGGGAGACCGCC 632

Qy 217 ValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 236
 Db 633 ATCCAGAAG-----CCCTGTGTGAGCCTC 656

Qy 237 ProAlaGlySerThrAspCysValCys-----TyrSerThr 248
 Db 657 CGAGCAGGCTCTGGCAACGGCTGGCAGCTTCCTTGNACCATATGCTGGCTATGACGAG 716

Qy 249 ValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla 268
 Db 717 GTCAACAAATGAAGACCTCTCTGACCAACTGCACGCTATTGCTGTGCCCGCGCTGCTGCA 776

Qy 269 ---MetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeu 287
 Db 777 CCATGAACCTGTGCTCTGCAACAGCGCTTCGGGGCTGCGCCTCTTCTCTGTGCTCAGC 836

Qy 288 LeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysValAsnTrpLeuGly 307

| | | | |
|----|------|---------------------------------------------------------------|------|
| Db | 837 | CTGGCCCTGGGGCTTCATTGCTGATGTGGACCTAGAGAGTGAGAAGTATCGGCGTCTGGGG | 896 |
| Qy | 308 | LeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGly | 327 |
| Db | 897 | GAGATGGCTTCACTCTGGGACATTCCTCGGTCTGGCAGCCCTGGCGACCTACCGCGC | 956 |
| Qy | 328 | ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCys | 347 |
| Db | 957 | CGATGGCCTACCTCCCTGTAGGAAGA--GTGGGTTCCAAGACACGTGCTCCCC-- | 1010 |
| Qy | 348 | ArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGluGlnLysLys | 367 |
| Db | 1011 | -----GTTGTGGTCCAGCAGGGCCCGGTAGATGCACAC----- | 1043 |
| Qy | 368 | AlaLeuTyrGlyLeuGluAlaAlaGluAspVal---GluGluTrpGlnValValCysGly | 386 |
| Db | 1044 | --CTTGTGCCACTGGAG-----GAGCCAGTGCCTCTCACTGGACAGTGGTGGCCGAC | 1094 |
| Qy | 387 | Lys-----PheLeuAlaAlaAsnAlaThrAsnMetSerCysAlaCysArgArg | 402 |
| Db | 1095 | GAGGACTTTGTGTAGTCTCTGGCACTCTGCATCGCACCTGGCGAGTGAAGTGTGGCT | 1154 |
| Qy | 403 | SerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeuLeu | 422 |
| Db | 1155 | GCACCCATGGCCGCTGTGCAGCT-----GGCGTCATGCATCTGTTCTAC | 1199 |
| Qy | 423 | IleArgLys---CysSerArgPheAsnPheLeuArg---PheLeuIleArgHisThrAsn | 440 |
| Db | 1200 | GTGGCGGGGAGTGTCTGCTGCCATGCTGCTGCCTCTTCTTGGCCATGGAGAAGGC | 1259 |
| Qy | 441 | GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe | 460 |
| Db | 1260 | AGGCATATGAGTATGAATGCCCTACTCTGGTATATGTGCCGTGGCTTCCGC--- | 1316 |
| Qy | 461 | ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLysLysArgPhe | 480 |
| Db | 1317 | -----TTGGAGCCCAAGAT-----GGAAAGGTGTGTTT | 1346 |
| Qy | 481 | GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn | 500 |
| Db | 1347 | GCA----- | 1349 |
| Qy | 501 | CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal | 520 |
| Db | 1350 | GTGATGGGGAATTGATGGTTAGCAGCGCGTGCAGGCCAGGTGCACCAAACTACTTC | 1409 |
| Qy | 521 | ArgPheAlaArgGlyIleGluGluAsnPro-----LysPro | 533 |
| Db | 1410 | TGGATGTTCAGCGGTGGTGGAGCCCGCCAGCTGGAAGGCC | 1454 |

Search completed: September 6, 2005, 02:34:14
Job time : 349.735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: September 3, 2005, 03:56:52 ; Search time 347.06 Seconds
(without alignments)
1891.374 Million cell updates/sec

Title: US-10-631-958-11
Perfect score: 3025
Sequence: 1 HEANGAPLGVAPPAPWT.....OLVRLFARGIENPKPDSHS 562

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
33: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
34: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
35: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
36: /cgn2_6/ptodata/1/paa/US110_COMB.pep.*
37: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|----------------------|
| 1 | 3025 | 100.0 | 562 | 25 | US-09-969-896-11 |
| 2 | 3025 | 100.0 | 562 | 32 | US-10-631-958-11 |
| 3 | 2888 | 95.5 | 537 | 1 | PCT-US02-24623-13 |
| 4 | 2888 | 95.5 | 537 | 1 | PCT-US02-31373-40 |
| 5 | 2888 | 95.5 | 537 | 1 | PCT-US03-21730-121 |
| 6 | 2888 | 95.5 | 537 | 25 | US-09-969-896-10 |
| 7 | 2888 | 95.5 | 537 | 28 | US-10-262-511A-40 |
| 8 | 2888 | 95.5 | 537 | 28 | US-10-262-511A-40 |
| 9 | 2888 | 95.5 | 537 | 30 | US-10-486-306-13 |
| 10 | 2888 | 95.5 | 537 | 32 | US-10-618-941-121 |
| 11 | 2888 | 95.5 | 537 | 32 | US-10-631-958-10 |
| 12 | 2880 | 95.2 | 537 | 29 | US-10-315-597A-2 |
| 13 | 2874.5 | 95.0 | 544 | 27 | US-10-170-205B-15652 |
| 14 | 2588 | 85.6 | 481 | 1 | PCT-US02-24623-12 |
| 15 | 2588 | 85.6 | 481 | 30 | US-10-486-306-12 |
| 16 | 2456.5 | 81.2 | 471 | 22 | US-09-784-810A-6 |
| 17 | 2456.5 | 81.2 | 471 | 22 | US-09-784-810C-6 |
| 18 | 2388.5 | 79.0 | 485 | 37 | US-60-243-468-1391 |
| 19 | 2350 | 77.7 | 536 | 1 | PCT-US03-30720-1291 |
| 20 | 2215 | 73.2 | 746 | 1 | PCT-US01-08631-43900 |
| 21 | 1714.5 | 56.7 | 727 | 1 | PCT-US01-08631-43902 |
| 22 | 1640.5 | 54.2 | 326 | 22 | US-09-784-810A-11 |
| 23 | 1640.5 | 54.2 | 326 | 22 | US-09-784-810C-11 |
| 24 | 1640.5 | 54.2 | 326 | 25 | US-09-969-896-2 |
| 25 | 1640.5 | 54.2 | 326 | 32 | US-10-631-958-2 |
| 26 | 1361.5 | 45.0 | 454 | 1 | PCT-US01-08631-43903 |
| 27 | 1159 | 38.3 | 255 | 1 | PCT-US03-30720-2504 |
| 28 | 1055 | 34.9 | 228 | 28 | US-10-218-140-3172 |
| 29 | 1032 | 34.1 | 190 | 1 | PCT-US01-42950-573 |
| 30 | 1032 | 34.1 | 190 | 30 | US-10-416-993-573 |
| 31 | 986 | 32.6 | 182 | 22 | US-09-784-810A-29 |
| 32 | 986 | 32.6 | 182 | 22 | US-09-784-810C-29 |
| 33 | 986 | 32.6 | 182 | 28 | US-10-218-140-4294 |
| 34 | 675 | 22.3 | 596 | 20 | US-09-614-150-35799 |
| 35 | 675 | 22.3 | 596 | 20 | US-09-614-150A-35799 |
| 36 | 675 | 22.3 | 596 | 37 | US-60-191-637-35410 |
| 37 | 664 | 22.0 | 136 | 1 | PCT-US01-08631-43899 |
| 38 | 647 | 21.4 | 144 | 22 | US-09-784-810A-8 |
| 39 | 647 | 21.4 | 144 | 22 | US-09-784-810C-8 |
| 40 | 628 | 20.8 | 136 | 1 | PCT-US01-08631-43901 |
| 41 | 611.5 | 20.2 | 588 | 30 | US-10-425-115-361076 |
| 42 | 597.5 | 19.8 | 575 | 30 | US-10-425-115-270874 |
| 43 | 594 | 19.6 | 532 | 37 | US-60-474-239-5 |
| 44 | 504.5 | 16.7 | 495 | 27 | US-10-170-205B-6540 |
| 45 | 480.5 | 15.9 | 613 | 30 | US-10-437-963-195284 |

ALIGNMENTS

RESULT 1
US-09-969-896-11
; Sequence 11, Application US/09969896
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 562
; TYPE: PRT

| | | | | | | | | | |
|--------------------------------------------------------------|-----|-----------------|--------------|--------------|-------------------------|---------------------|------------|-----|--|
| ; ORGANISM: Homo sapiens | | | | | | | | | |
| US-09-969-896-11 | | | | | | | | | |
| Query Match 100.0%; Score 3025; DB 25; Length 562; | | | | | | | | | |
| Best Local Similarity 100.0%; Pred. No. 1e-292; | | | | | | | | | |
| Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| Qy | 1 | HEAANGPAPLGVAPP | PAWRTSPA | EMGATGAA | EPLQSVLWVKQORCAV | SVLEPARALLRW | 60 | | |
| Db | 1 | HEAANGPAPLGVAPP | PAWRTSPA | EMGATGAA | EPLQSVLWVKQORCAV | SVLEPARALLRW | 60 | | |
| Qy | 61 | RSPGAGAGACADAC | SPVSEII | IAVEETDVHGK | HQSGKQKMEKPYAFTV | HCVKRRRH | 120 | | |
| Db | 61 | RSPGAGAGACADAC | SPVSEII | IAVEETDVHGK | HQSGKQKMEKPYAFTV | HCVKRRRH | 120 | | |
| Qy | 121 | RWKQAQVTFWCP | BEQCHLWLT | REMLEKLT | SRPKHLLVF | INPFGKGQKRIYERKVA | 180 | | |
| Db | 121 | RWKQAQVTFWCP | BEQCHLWLT | REMLEKLT | SRPKHLLVF | INPFGKGQKRIYERKVA | 180 | | |
| Qy | 181 | PLFTLASITTTDII | IVTEHANQAKET | LYEINIDKYDGI | VCVGGDMFSEVLHGLIGRTORSA | 240 | | | |
| Db | 181 | PLFTLASITTTDII | IVTEHANQAKET | LYEINIDKYDGI | VCVGGDMFSEVLHGLIGRTORSA | 240 | | | |
| Qy | 241 | GVQNHPRAVLVP | SSLRIGIIP | PAGSTDVCY | STVGTSDAET | SALHIVVGD | SLAMDVSSVH | 300 | |
| Db | 241 | GVQNHPRAVLVP | SSLRIGIIP | PAGSTDVCY | STVGTSDAET | SALHIVVGD | SLAMDVSSVH | 300 | |
| Qy | 301 | HNSTLLRYSVLL | GYFGYDII | IKDSEKRWLG | LARYDFSLK | TFLSHHCYEGTVSFLPAQ | 360 | | |
| Db | 301 | HNSTLLRYSVLL | GYFGYDII | IKDSEKRWLG | LARYDFSLK | TFLSHHCYEGTVSFLPAQ | 360 | | |
| Qy | 361 | HTVGSPRDRKPC | RACGFCV | CRSQKQLE | EEQKALYGLEAAEDVE | EWQVCGKFLAINATN | 420 | | |
| Db | 361 | HTVGSPRDRKPC | RACGFCV | CRSQKQLE | EEQKALYGLEAAEDVE | EWQVCGKFLAINATN | 420 | | |
| Qy | 421 | MSCACRRSPRGL | SPAHLG | DGSSDLIL | IRKCSRPNFLR | FLIRHTNQDQDFTFVEYRV | 480 | | |
| Db | 421 | MSCACRRSPRGL | SPAHLG | DGSSDLIL | IRKCSRPNFLR | FLIRHTNQDQDFTFVEYRV | 480 | | |
| Qy | 481 | KKFOFTSKHME | DESDLKEG | KKRFGHIC | SSHPSCCCTV | SNWNCDEVLHSPAIEVRV | 540 | | |
| Db | 481 | KKFOFTSKHME | DESDLKEG | KKRFGHIC | SSHPSCCCTV | SNWNCDEVLHSPAIEVRV | 540 | | |
| Qy | 541 | HCQLVRLFARG | IEENPK | PDSSH | 562 | | | | |
| Db | 541 | HCQLVRLFARG | IEENPK | PDSSH | 562 | | | | |
| RESULT 2 | | | | | | | | | |
| US-10-631-958-11 | | | | | | | | | |
| ; Sequence 11, Application US/10631958 | | | | | | | | | |
| ; GENERAL INFORMATION: | | | | | | | | | |
| ; APPLICANT: Kossida, Sophia | | | | | | | | | |
| ; TITLE OF INVENTION: Regulation of human Sphingosine | | | | | | | | | |
| ; FILE REFERENCE: 004974.00594 | | | | | | | | | |
| ; CURRENT APPLICATION NUMBER: US/10/631,958 | | | | | | | | | |
| ; PRIOR FILING DATE: 2003-08-01 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US/09/969,896 | | | | | | | | | |
| ; PRIOR FILING DATE: 2001-10-04 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/238,005 | | | | | | | | | |
| ; PRIOR FILING DATE: 2000-10-06 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/314,113 | | | | | | | | | |
| ; PRIOR FILING DATE: 2001-08-23 | | | | | | | | | |
| ; NUMBER OF SEQ ID NOS: 16 | | | | | | | | | |
| ; SOFTWARE: FastSeq for Windows Version 4.0 | | | | | | | | | |
| ; SEQ ID NO 11 | | | | | | | | | |
| ; LENGTH: 562 | | | | | | | | | |
| ; TYPE: PRT | | | | | | | | | |
| ; ORGANISM: Homo sapiens | | | | | | | | | |
| US-10-631-958-11 | | | | | | | | | |
| Query Match 100.0%; Score 3025; DB 32; Length 562; | | | | | | | | | |

| | | | | | | | | | |
|-------------------------------------------------------------------------------|-----|-----------------|--------------|--------------|-------------------------|---------------------|------------|-----|--|
| Best Local Similarity 100.0%; Pred. No. 1e-292; | | | | | | | | | |
| Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| Qy | 1 | HEAANGPAPLGVAPP | PAWRTSPA | EMGATGAA | EPLQSVLWVKQORCAV | SVLEPARALLRW | 60 | | |
| Db | 1 | HEAANGPAPLGVAPP | PAWRTSPA | EMGATGAA | EPLQSVLWVKQORCAV | SVLEPARALLRW | 60 | | |
| Qy | 61 | RSPGAGAGACADAC | SPVSEII | IAVEETDVHGK | HQSGKQKMEKPYAFTV | HCVKRRRH | 120 | | |
| Db | 61 | RSPGAGAGACADAC | SPVSEII | IAVEETDVHGK | HQSGKQKMEKPYAFTV | HCVKRRRH | 120 | | |
| Qy | 121 | RWKQAQVTFWCP | BEQCHLWLT | REMLEKLT | SRPKHLLVF | INPFGKGQKRIYERKVA | 180 | | |
| Db | 121 | RWKQAQVTFWCP | BEQCHLWLT | REMLEKLT | SRPKHLLVF | INPFGKGQKRIYERKVA | 180 | | |
| Qy | 181 | PLFTLASITTTDII | IVTEHANQAKET | LYEINIDKYDGI | VCVGGDMFSEVLHGLIGRTORSA | 240 | | | |
| Db | 181 | PLFTLASITTTDII | IVTEHANQAKET | LYEINIDKYDGI | VCVGGDMFSEVLHGLIGRTORSA | 240 | | | |
| Qy | 241 | GVQNHPRAVLVP | SSLRIGIIP | PAGSTDVCY | STVGTSDAET | SALHIVVGD | SLAMDVSSVH | 300 | |
| Db | 241 | GVQNHPRAVLVP | SSLRIGIIP | PAGSTDVCY | STVGTSDAET | SALHIVVGD | SLAMDVSSVH | 300 | |
| Qy | 301 | HNSTLLRYSVLL | GYFGYDII | IKDSEKRWLG | LARYDFSLK | TFLSHHCYEGTVSFLPAQ | 360 | | |
| Db | 301 | HNSTLLRYSVLL | GYFGYDII | IKDSEKRWLG | LARYDFSLK | TFLSHHCYEGTVSFLPAQ | 360 | | |
| Qy | 361 | HTVGSPRDRKPC | RACGFCV | CRSQKQLE | EEQKALYGLEAAEDVE | EWQVCGKFLAINATN | 420 | | |
| Db | 361 | HTVGSPRDRKPC | RACGFCV | CRSQKQLE | EEQKALYGLEAAEDVE | EWQVCGKFLAINATN | 420 | | |
| Qy | 421 | MSCACRRSPRGL | SPAHLG | DGSSDLIL | IRKCSRPNFLR | FLIRHTNQDQDFTFVEYRV | 480 | | |
| Db | 421 | MSCACRRSPRGL | SPAHLG | DGSSDLIL | IRKCSRPNFLR | FLIRHTNQDQDFTFVEYRV | 480 | | |
| Qy | 481 | KKFOFTSKHME | DESDLKEG | KKRFGHIC | SSHPSCCCTV | SNWNCDEVLHSPAIEVRV | 540 | | |
| Db | 481 | KKFOFTSKHME | DESDLKEG | KKRFGHIC | SSHPSCCCTV | SNWNCDEVLHSPAIEVRV | 540 | | |
| Qy | 541 | HCQLVRLFARG | IEENPK | PDSSH | 562 | | | | |
| Db | 541 | HCQLVRLFARG | IEENPK | PDSSH | 562 | | | | |
| RESULT 3 | | | | | | | | | |
| PCT-US02-24623-13 | | | | | | | | | |
| ; Sequence 13, Application PC/TUS0224623 | | | | | | | | | |
| ; GENERAL INFORMATION: | | | | | | | | | |
| ; APPLICANT: EXELIXIS, INC. | | | | | | | | | |
| ; TITLE OF INVENTION: SPKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE | | | | | | | | | |
| ; FILE REFERENCE: EX02-104 | | | | | | | | | |
| ; CURRENT APPLICATION NUMBER: PCT/US02/24623 | | | | | | | | | |
| ; CURRENT FILING DATE: 2002-08-02 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/310,362 | | | | | | | | | |
| ; PRIOR FILING DATE: 2001-08-06 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/357,501 | | | | | | | | | |
| ; PRIOR FILING DATE: 2002-02-15 | | | | | | | | | |
| ; NUMBER OF SEQ ID NOS: 13 | | | | | | | | | |
| ; SOFTWARE: Patent in version 3.1 | | | | | | | | | |
| ; SEQ ID NO 13 | | | | | | | | | |
| ; LENGTH: 537 | | | | | | | | | |
| ; TYPE: PRT | | | | | | | | | |
| ; ORGANISM: Homo sapiens | | | | | | | | | |
| PCT-US02-24623-13 | | | | | | | | | |
| Query Match 95.5%; Score 2888; DB 1; Length 537; | | | | | | | | | |
| Best Local Similarity 100.0%; Pred. No. 5e-279; | | | | | | | | | |
| Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| Qy | 26 | MGATGAAPLOS | VLWVKQORCAV | SVLEPARALLRW | SPGAGAGAC | AGADACSV | PSVSEII | 85 | |
| Db | 1 | MGATGAAPLOS | VLWVKQORCAV | SVLEPARALLRW | SPGAGAGAC | AGADACSV | PSVSEII | 60 | |

QY 86 BETDVHGKHGGSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQCHLWLQTLR 145
Db 61 ESTDVHGKHGGSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQCHLWLQTLR 120
QY 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAPLFTLASITTDIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAPLFTLASITTDIIVTEHANOAKETL 180
QY 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 240
QY 266 TDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGIFYGDI I KDS 325
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGIFYGDI I KDS 300
QY 326 EKRWGLGARYDFSGLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 385
Db 301 EKRWGLGARYDFSGLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 360
QY 386 LEEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 445
Db 361 LEEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
QY 446 ILIRKCSRFNRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKGGKKRF 505
Db 421 ILIRKCSRFNRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKGGKKRF 480
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 562
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537

RESULT 4

PCT-US02-31373-40

; Sequence 40, Application PC/TUS0231373

; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation

; APPLICANT: Smithson, Glenda

; APPLICANT: Millet, Isabelle

; APPLICANT: Peyman, John A.

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Ju, Jingfang

; APPLICANT: Li, Li

; APPLICANT: Guo, Xiaojia (Sasha)

; APPLICANT: Patturajan, Meera

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Ellerman, Karen

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Ort, Tatiana

; APPLICANT: Gorman, Linda

; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Anderson, David W.

; APPLICANT: Zhong, Mei

; APPLICANT: Catterton, Elina

; APPLICANT: Ji, Weizhen

; APPLICANT: Miller, Charles E.

; APPLICANT: Rastelli, Luca

; APPLICANT: Stone, David J.

; APPLICANT: Pena, Carol E. A.

; APPLICANT: Shenoy, Suresh G.

; APPLICANT: Shinkets, Richard A.

; APPLICANT: Rothenberg, Mark E.

; APPLICANT: Leach, Martin D.

; APPLICANT: Ages, Michele L.

; APPLICANT: Berghs, Constance

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 21402-462C-061

; CURRENT APPLICATION NUMBER: PCT/US02/31373

; CURRENT FILING DATE: 2003-09-28

; PRIOR APPLICATION NUMBER: 10/262,511

; PRIOR FILING DATE: 2002-10-01

; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqlist version 0.1
; SEQ ID NO 40
; LENGTH: 537
; TYPE: PRT
; ORGANISM: CG143216-01
PCT-US02-31373-40

Query Match 95.5%; Score 2888; DB 1; Length 537;

Best Local Similarity 100.0%; Pred. No. 5e-279;

Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 MGATGAAPLQSLVWVQQRCAVSLPARALLRWRSPGPGAGAPGADACSVPVSEIIAV 85
Db 1 MGATGAAPLQSLVWVQQRCAVSLPARALLRWRSPGPGAGAPGADACSVPVSEIIAV 60
QY 86 BETDVHGKHGGSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQCHLWLQTLR 145
Db 61 ESTDVHGKHGGSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQCHLWLQTLR 120
QY 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAPLFTLASITTDIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAPLFTLASITTDIIVTEHANOAKETL 180
QY 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 240
QY 266 TDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGIFYGDI I KDS 325
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGIFYGDI I KDS 300
QY 326 EKRWGLGARYDFSGLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 385
Db 301 EKRWGLGARYDFSGLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 360
QY 386 LEEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 445
Db 361 LEEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
QY 446 ILIRKCSRFNRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKGGKKRF 505
Db 421 ILIRKCSRFNRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKGGKKRF 480
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 562
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537

RESULT 5

PCT-US03-21730-121

; Sequence 121, Application PC/TUS0321730

; GENERAL INFORMATION:

; APPLICANT: SUGEN, INC.

```
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536/0461
; CURRENT APPLICATION NUMBER: PCT/US03/21730
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-21730-121

Query Match          95.5%; Score 2888; DB 1; Length 537;
Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAPLOSVLWKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 85
Db 1 MGATGAAPLOSVLWKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60

Qy 86 EETDVHGKHQSGKWQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 145
Db 61 EETDVHGKHQSGKWQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 120

Qy 146 EMLEKLTSRPKHLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANOAKETL 180

Qy 206 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 240

Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 325
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 300

Qy 326 EKKRWLGLARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 385
Db 301 EKKRWLGLARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 360

Qy 386 LEEBQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 445
Db 361 LEEBQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420

Qy 446 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGKKRF 505
Db 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGKKRF 480

Qy 506 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 562
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 537

RESULT 6
US-09-969-896-10
; Sequence 10, Application US/09969896
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennada
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-896-10

Query Match          95.5%; Score 2888; DB 25; Length 537;
Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAPLOSVLWKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 85
Db 1 MGATGAAPLOSVLWKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60

Qy 86 EETDVHGKHQSGKWQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 145
Db 61 EETDVHGKHQSGKWQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR 120

Qy 146 EMLEKLTSRPKHLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANOAKETL 180

Qy 206 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 240

Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 325
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 300

Qy 326 EKKRWLGLARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 385
Db 301 EKKRWLGLARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 360

Qy 386 LEEBQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 445
Db 361 LEEBQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420

Qy 446 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGKKRF 505
Db 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGKKRF 480

Qy 506 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 562
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 537

RESULT 7
US-10-262-511-40
; Sequence 40, Application US/10262511
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennada
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
```



```
Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAEPLOSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 85
Db 1 MGATGAAEPLOSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 60
Qy 86 BETDVHGKHQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQI.CHLWLOTLR 145
Db 61 BETDVHGKHQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQI.CHLWLOTLR 120
Qy 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 206 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIDKDS 325
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIDKDS 300
Qy 326 EKRWLGLARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRACGFCVCRQSKQQ 385
Db 301 EKRWLGLARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRACGFCVCRQSKQQ 360
Qy 386 LEBEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445
Db 361 LEBEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Qy 446 ILIRKCSRFNRLFLIRHTNQDQDFTFVVEYRVKKFOFTSKHMEDESDLKEGGKKRF 505
Db 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVVEYRVKKFOFTSKHMEDESDLKEGGKKRF 480
Qy 506 GHICSSHPPCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSSH 562
Db 481 GHICSSHPPCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSSH 537

RESULT 9
US-10-486-306-13
; Sequence 13, Application US/10486306
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SPKS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-104C-US
; CURRENT APPLICATION NUMBER: US/10/486,306
; PRIOR FILING DATE: 2004-02-05
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/310,362
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/357,501
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-486-306-13

Query Match 95.5%; Score 2888; DB 30; Length 537;
Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAEPLOSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 85
Db 1 MGATGAAEPLOSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 60
Qy 86 BETDVHGKHQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQI.CHLWLOTLR 145
Db 61 BETDVHGKHQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQI.CHLWLOTLR 120
Qy 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 206 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIDKDS 325
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIDKDS 300

Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAEPLOSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 85
Db 1 MGATGAAEPLOSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 60
Qy 86 BETDVHGKHQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQI.CHLWLOTLR 145
Db 61 BETDVHGKHQSGKWQKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQI.CHLWLOTLR 120
Qy 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 206 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIDKDS 325
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIDKDS 300
```

QY 326 EKKRWGLGARYDFSLGKTLFSLHHCEYGVTSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 385
DB 301 EKKRWGLGARYDFSLGKTLFSLHHCEYGVTSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 360
QY 386 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445
DB 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
QY 446 ILIRKCSRNFRLFLIRHTNQDQDFTFVEVYRVKPKQFTSKHMEDESDSKEGGKKRF 505
DB 421 ILIRKCSRNFRLFLIRHTNQDQDFTFVEVYRVKPKQFTSKHMEDESDSKEGGKKRF 480
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 562
DB 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537

RESULT 11

US-10-631-958-10
; Sequence 10, Application US/10631958
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-958-10

Query Match 95.5%; Score 2888; DB 32; Length 537;
Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 MGATGAAPLQSVLVVKQORCAVSLPARALLRWRSPGPGAGAPGADACSVPVSEIIAV 85
DB 1 MGATGAAPLQSVLVVKQORCAVSLPARALLRWRSPGPGAGAPGADACSVPVSEIIAV 60
QY 86 EETDVHGKHQSGKWKQMEKPYAFTVHCVRARRHRWKWAQVTFWCPPEQLCHLWLTLR 145
DB 61 EETDVHGKHQSGKWKQMEKPYAFTVHCVRARRHRWKWAQVTFWCPPEQLCHLWLTLR 120
QY 146 EMLEKLTSRPKHLVFINPFGKGQGRKRIYERKVAPLFTLASITTDIIIVTEHANOAKETL 205
DB 121 EMLEKLTSRPKHLVFINPFGKGQGRKRIYERKVAPLFTLASITTDIIIVTEHANOAKETL 180
QY 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 265
DB 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 240
QY 266 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFGYGDIIKDS 325
DB 241 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFGYGDIIKDS 300
QY 326 EKKRWGLGARYDFSLGKTLFSLHHCEYGVTSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 385
DB 301 EKKRWGLGARYDFSLGKTLFSLHHCEYGVTSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 360
QY 386 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445
DB 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
QY 446 ILIRKCSRNFRLFLIRHTNQDQDFTFVEVYRVKPKQFTSKHMEDESDSKEGGKKRF 505
DB 421 ILIRKCSRNFRLFLIRHTNQDQDFTFVEVYRVKPKQFTSKHMEDESDSKEGGKKRF 480
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 562
DB 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537

QY 446 ILIRKCSRNFRLFLIRHTNQDQDFTFVEVYRVKPKQFTSKHMEDESDSKEGGKKRF 505
DB 421 ILIRKCSRNFRLFLIRHTNQDQDFTFVEVYRVKPKQFTSKHMEDESDSKEGGKKRF 480
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 562
DB 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537

RESULT 12

US-10-315-597A-2
; Sequence 2, Application US/10315597A
; GENERAL INFORMATION:
; APPLICANT: Sugiyura, Masako
; APPLICANT: Kono, Keita
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It
; FILE REFERENCE: 02658CIP/HG
; CURRENT APPLICATION NUMBER: US/10/315,597A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: JP 2000-178039
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-315-597A-2

Query Match 95.2%; Score 2880; DB 29; Length 537;
Best Local Similarity 99.6%; Pred. No. 3.2e-278;
Matches 535; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 26 MGATGAAPLQSVLVVKQORCAVSLPARALLRWRSPGPGAGAPGADACSVPVSEIIAV 85
DB 1 MGATGAAPLQSVLVVKQORCAVSLPARALLRWRSPGPGAGAPGADACSVPVSEIIAV 60
QY 86 EETDVHGKHQSGKWKQMEKPYAFTVHCVRARRHRWKWAQVTFWCPPEQLCHLWLTLR 145
DB 61 EETDVHGKHQSGKWKQMEKPYAFTVHCVRARRHRWKWAQVTFWCPPEQLCHLWLTLR 120
QY 146 EMLEKLTSRPKHLVFINPFGKGQGRKRIYERKVAPLFTLASITTDIIIVTEHANOAKETL 205
DB 121 EMLEKLTSRPKHLVFINPFGKGQGRKRIYERKVAPLFTLASITTDIIIVTEHANOAKETL 180
QY 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 265
DB 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 240
QY 266 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFGYGDIIKDS 325
DB 241 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFGYGDIIKDS 300
QY 326 EKKRWGLGARYDFSLGKTLFSLHHCEYGVTSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 385
DB 301 EKKRWGLGARYDFSLGKTLFSLHHCEYGVTSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 360
QY 386 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445
DB 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
QY 446 ILIRKCSRNFRLFLIRHTNQDQDFTFVEVYRVKPKQFTSKHMEDESDSKEGGKKRF 505
DB 421 ILIRKCSRNFRLFLIRHTNQDQDFTFVEVYRVKPKQFTSKHMEDESDSKEGGKKRF 480
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 562
DB 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537

RESULT 13

```

US-10-170-205E-15652
; Sequence 15652, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15652
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-15652

Query Match          95.0%; Score 2874.5; DB 27; Length 544;
Best Local Similarity 98.7%; Pred. No. 1.2e-277;
Matches 537; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

QY      26  MGATGAEPQLQSVLVWKQQRCAVSLPEPARALLRWRSPPGAGAGPGA-----DACSVP 78
DB      1  MGATGAEPQLQSVLVWKQQRCAVSLPEPARALLRWRSPPGAGAGPGA-----DACSVP 60

QY      79  VSEIIAEEETDVHGKHQSGKQKMEKPYAFTVHCAGARRHRWKAQVTFWCPEEQQLCH 138
DB      61  VSEIIAEEETDVHGKHQSGKQKMEKPYAFTVHCAGARRHRWKAQVTFWCPEEQQLCH 120

QY      139 LWLQTLREMLEKUTSRPKHLLVFINPFGKGQGGKIYERKVAPLFTLASITTDIIIVTEHA 198
DB      121 LWLQTLREMLEKUTSRPKHLLVFINPFGKGQGGKIYERKVAPLFTLASITTDIIIVTEHA 180

QY      199 NQAKETLYEINIDKYGIVCVGDDGMFSEVLHGLTGRTORSAGVDQNHPRAVLVPSSLRI 258
DB      181 NQAKETLYEINIDKYGIVCVGDDGMFSEVLHGLTGRTORSAGVDQNHPRAVLVPSSLRI 240

QY      259 GIIPAGSTDCVCYSTVTGTSDAETSALHIVVGDSLAMDYSSVHHNSTLLRYSVSLGIGFY 318
DB      241 GIIPAGSTDCVCYSTVTGTSDAETSALHIVVGDSLAMDYSSVHHNSTLLRYSVSLGIGFY 300

QY      319 GDIIKDKSEKRWLGLARYDFSGLKTPLSHHCYEGTVTSFLPAQHTVGSPRDRKPCRAGCFV 378
DB      301 GDIIKDKSEKRWLGLARYDFSGLKTPLSHHCYEGTVTSFLPAQHTVGSPRDRKPCRAGCFV 360

QY      379 CROSKOOLEEOKKALYGLEAAEDVEEQVCGKFLAATNMTWSCACRRSPRGLSPAHL 438
DB      361 CROSKOOLEEOKKALYGLEAAEDVEEQVCGKFLAATNMTWSCACRRSPRGLSPAHL 420

QY      439 GDGSSDLLIRKCSRFNRLFLIRHNTQQDFTFVEYVRVKKQFTSKMEDESDLK 498
DB      421 GDGSSDLLIRKCSRFNRLFLIRHNTQQDFTFVEYVRVKKQFTSKMEDESDLK 480

QY      499 EGCKKRGFHTCSHSPSCCTTNSNSNCDGEVLHSPAIEVRVHCOLVRLFARGIENPKP 558
DB      481 EGCKKRGFHTCSHSPSCCTTNSNSNCDGEVLHSPAIEVRVHCOLVRLFARGIENPKP 540

QY      559 DSHS 562
DB      541 DSHS 544

RESULT 14
PCT-US02-24623-12
; Sequence 12, Application PC/TUS0224623
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SPHKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-104
; CURRENT APPLICATION NUMBER: PCT/US02/24623
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/310,362
; PRIOR FILING DATE: 2001-08-06

```

| | | | |
|---------------------------------------------------|------|----------------------------------------------------------------|---------------------------------|
| Best Local Similarity 100.0%, Pred. No. 4.8e-249; | | | |
| Matches | 481; | Conservative 0; | Mismatches 0; Indels 0; Gaps 0; |
| Qy | 82 | IIAVEETDVHGKHGQSGKWQKMEKPYAFVTHCVKRRHRHWKQAQVTFWCPEEQ.LCHLWL | 141 |
| Db | 1 | IIAVEETDVHGKHGQSGKWQKMEKPYAFVTHCVKRRHRHWKQAQVTFWCPEEQ.LCHLWL | 60 |
| Qy | 142 | QTLREMLEKLTSRPKHLLVFNPFGGKGQGRRIYVERKVAPLFTTLASITTD.IIVTEHANQA | 201 |
| Db | 61 | QTLREMLEKLTSRPKHLLVFNPFGGKGQGRRIYVERKVAPLFTTLASITTD.IIVTEHANQA | 122 |
| Qy | 202 | KETIYEINIDKYLVCVGGDMSEVLHGLIGRTQRSAGVDQNHPRALVPSSLRIGII | 261 |
| Db | 121 | KETIYEINIDKYLVCVGGDMSEVLHGLIGRTQRSAGVDQNHPRALVPSSLRIGII | 180 |
| Qy | 262 | PAGSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLLGXYGYGDI | 321 |
| Db | 181 | PAGSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLLGXYGYGDI | 240 |
| Qy | 322 | IKDSEKKRWGLGARYDFSLGKATFLSHHCYEGTVSFLPAQHTVVGSPDRKKPCRACFCVCRQ | 381 |
| Db | 241 | IKDSEKKRWGLGARYDFSLGKATFLSHHCYEGTVSFLPAQHTVVGSPDRKKPCRACFCVCRQ | 300 |
| Qy | 382 | SKQOLEEBEQKALYGLEAAEDVEEVQVCGKFLAINATNMSCACRRSPRGLSPAHLGDG | 441 |
| Db | 301 | SKQOLEEBEQKALYGLEAAEDVEEVQVCGKFLAINATNMSCACRRSPRGLSPAHLGDG | 360 |
| Qy | 442 | SSDILIRKCSRFPNPLRIBHNTQQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGG | 501 |
| Db | 361 | SSDILIRKCSRFPNPLRIBHNTQQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGG | 420 |
| Qy | 502 | KKRFGHICSSHPPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGTEENPKPDH | 561 |
| Db | 421 | KKRFGHICSSHPPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGTEENPKPDH | 480 |
| Qy | 562 | S 562 | |
| Db | 481 | S 481 | |

Search completed: September 3, 2005, 04:25:25
Job time : 349.06 secs

This Page Blank (uspto)